PROSUMER-BASED DECENTRALIZED UNIT COMMITMENT FOR FUTURE ELECTRICITY GRIDS

A Thesis
Presented to
The Academic Faculty

by

Mitcham Costley

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2015

PROSUMER-BASED DECENTRALIZED UNIT COMMITMENT FOR FUTURE ELECTRICITY GRIDS

Approved by:

Professor Santiago Grijalva, Advisor
Associate Professor, School of ECE
Georgia Institute of Technology

Professor A. P. Meliopoulos
Professor, School of ECE
Georgia Institute of Technology

Professor Magnus Egerstedt
Professor, School of ECE
Georgia Institute of Technology

Professor Bonnie Ferri
Professor, School of ECE
Georgia Institute of Technology

Professor Shabbir Ahmed
Professor, School of ISyE
Georgia Institute of Technology

Professor Maryam Saeedifard
Assistant Professor, School of ECE
Georgia Institute of Technology

Date Approved: March 30, 2015

To God, to whom all praise is due.

To James and Cecelia Costley, for their unflagging support of me in this work. I am forever indebted to them and my family.

To Betsy White, without whom the last part of this process would have been far more difficult.

ACKNOWLEDGEMENTS

I consider the opportunities I have been given in the Ph.D. program at Georgia Tech to be among the most challenging and formative of my life. I have grown personally and professionally in ways that I did not expect. I have been a member on teams whose diversity of background and technical expertise surprised and humbled me. I was able to make a small contribution to the state of engineering knowledge. The people who surrounded me did not just help me through this experience; to a large extent, they *were* the experience.

I certainly want to thank my research advisor, Dr. Santiago Grijalva, for his continuous support. As my advisor, he helped me discover my strengths and interests, guiding me toward research questions that I could help answer. Moreover, he gave me many unexpected opportunities for growth, from international travel to high-level research presentations. Today I feel well-equipped to not only perform technical work but also communicate the importance and the results of it. I owe a large part of this growth to Dr. Grijalva. Also, he has always advocated externally for my work, building interest and support for it so that I could focus on the research effort. This too was very helpful to me and deserving of many thanks.

To Drs. Magnus Egerstedt and Shabbir Ahmed I also owe thanks for their consistent guidance in research direction during the ARPA-E GENI project on which I spent much of my time. Their mastery of their technical areas was invaluable when my research extended outside my formal background. Special thanks also go to Dr. Sakis Meliopoulos, Dr. Maryam Saeedifard, and Dr. Bonnie Ferri for serving on my defense exam committee. Their time investment is very much appreciated.

I have a deep debt of gratitude to Mohammad Javad Feizollahi for being such

an excellent colleague and research partner. Working with him has been one of the best collaborative experiences of my professional career. I am very grateful to have had the chance to learn from him, not only about optimization theory and operations research, but also about the poetry of Rumi and the history of Persia.

Of course, I must thank everyone in the ACES lab for being wonderful labmates, colleagues, and friends. It has been a deep pleasure to work with all of them and learn from them. The ACES lab is noteworthy for the breadth of experience that is represented, and I feel privileged to have been a part of it. I especially want to thank Alyse Taylor and Tanguy Hubert, who since the earliest days of the lab have been a positive presence and source of support to me. I don't know whether I would have finished this program if they weren't here, but I do know it would have been less fun. I also must thank ACES alumni Nathan Ainsworth for his leadership early in my program and Taylor Hollis for our valuable discussions on the application of research results in the field.

I also want to thank Dr. Whit Smith for the mentoring he has provided to me over the last 15 years. Ever since my early research projects in high school, he has helped me understand how to apply a critical mindset to anticipate and solve problems. He has taught me more than he knows, and I consider him one of my chief role models as an engineer.

Thanks to Glen Stewart and Frederick Ealick, who contributed greatly to the accomplishment of this work by being excellent and understanding roommates and great friends. Also thanks to all of my close friends from my undergraduate days at Georgia Tech who encouraged me to pursue this challenge.

Special thanks to the Department of Energy and ARPA-E for funding the research effort that led to this dissertation. It would not have been possible without their significant investment in me and in our team. The level of resources they made available to us is rare and is deeply appreciated.

Finally, I must thank my family at North Avenue Presbyterian Church for welcoming, teaching, supporting, and challenging me the last 5 years. I have many to thank there: Brooke and Michael Beckert for their longstanding friendship, Jeff Meyers and Courtney Quiros for teaching of all kinds, Michael Payne for deep conversations, Eddie Turner for action items, and too many others to name. They have had an impact I can never repay.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The contributions of this research are a scalable formulation and solution method for decentralized unit commitment, experimental results comparing decentralized unit commitment solution times to conventional unit commitment methods, a demonstration of the benefits of faster unit commitment computation time, and extensions of decentralized unit commitment to handle system network security constraints. We begin with a discussion motivating the shift from centralized power system control architectures to decentralized architectures and describe the characteristics of such an architecture. We then develop a formulation and solution method to solve decentralized unit commitment by adapting an existing approach for separable convex optimization problems to the nonconvex domain of unit commitment. The potential computational speed benefits of the novel decentralized unit commitment approach are then further investigated through a rolling-horizon framework that represents how system operators make decisions and adjustments online as new information is revealed. Finally, the decentralized unit commitment approach is extended to include network contingency constraints, a crucial function for the maintenance of system security. The results indicate decentralized unit commitment holds promise as a way of coordinating system operations in a future decentralized grid and also may provide a way to leverage parallel computing resources to solve large-scale unit commitment problems with greater speed and model fidelity than is possible with conventional methods.

CHAPTER I

INTRODUCTION

Software applications for power system management include several high-level functions, the most critical among them state estimation, unit commitment, economic dispatch, and security assessment and contingency analysis. In this work, we will concern ourselves primarily with the unit commitment problem, which possesses some overlap with the economic dispatch problem. Simulations capturing the interactions between unit commitment and economic dispatch will lend perspective to the importance of having a well-designed operations framework. However, all novel contributions center on unit commitment problems.

In this work, the electricity industry evolution toward decentralized power system architectures is anticipated. A decentralized framework requires the definition of some enclosed boundaries in which is encapsulated some information not directly accessible outside the grid except through specific functionality. The agents that represent these regions of the global system are called "prosumers." Each prosumer may consume, produce, store, or transport electricity among the others in the network depending on the functionality and capabilities they advertise. The prosumer is envisioned as the vehicle through which data is encapsulated and complexity is hidden to simplify the large-scale information infrastructures that will be needed in the future grid.

Nonetheless, the functionality represented by traditional grid operations remains essential. Methods for solving the coordination, scheduling, and control problems of today must be available and compatible with the computationally distributed and structurally decentralized model of the future grid. In this work, an algorithm is developed that addresses the need for decentralized unit commitment. The diversity

of time scales over which unit commitment problems are conducted, from weeks and months ahead to the sub-hour level, means that solving unit commitment in a decentralized way provides a strong foundation for the general time-coupled coordination of interconnected prosumers. Unit commitment also spans several organizational layers from electricity markets (satisfying economic objectives) to system control (satisfying reliability objectives). The breadth of potential impact for decentralized unit commitment is therefore wide for future grid control.

However, we also recognize that fully decentralized grid operations remain many years away from broad implementation. Therefore, the observed potential for benefits in computational speed are also emphasized as a more short-term benefit for system operators. These benefits can be expressed in several forms. Unit commitment problems, which are expanding in size as markets grow due to consolidation and the increase of virtual bidding, are becoming more difficult to solve. Any method that can help maintain the ability to find good solutions in the required time frames can help. Additionally, problems can be solved with greater model fidelity. For example, system operators typically only monitor the capacity limits of transmission lines that are expected to be congested. If the decentralized approach allows more transmission lines constraints to be included, the decision as to which line constraints need to be included becomes less critical. Further gains may be seen in the coordination of transactions between systems. Topology models are often not synchronized between systems, and a decentralized approach can speed up the process of reviewing exchanges by not requiring the consolidation of a global system model in a single mainframe.

It should be emphasized that all performance comparisons of decentralized unit commitment methods to conventional centralized unit commitment methods are made based on the assumption that the conventional unit commitment formulation and software implementation are representative of the current state of the art. Although care

has been taken to ensure the most recent literature on high-performing unit commitment formulations has been incorporated into our centralized model to provide a reasonable baseline, commercial unit commitment engines include many proprietary techniques to increase solution quality and decrease solution time. Therefore, rather than claiming that the decentralized unit commitment methods presented here solve large-scale problems faster than conventional methods in general practice or centralized methods in theory, we instead provide these results to indicate that performance of the decentralized unit commitment is within reasonable tolerances for solution time and solution quality. One of the primary contributions of this work, then, is the demonstration of feasibility of solving unit commitment problems in a decentralized way.

Focusing primarily on the latter category of applications, we develop a framework that helps to quantify the benefit from solving unit commitment with greater speed and with faster periodicity. In real control centers, unit commitment is solved repeatedly on a "rolling-horizon" basis to make adjustments to device commands based on new information about load, generation availability, and other aspects of system state. This rolling-horizon process is rarely addressed in the literature even though it is the reality of system operations. We discuss and demonstrate the mechanisms by which faster unit commitment can lend greater reliability and reduced cost to the grid while contributing formulations for these subproblems not available in the academic literature.

We also present two methods for extending decentralized unit commitment to the contingency-constrained case to enforce system network security. Network contingency constraints are crucial to satisfy required reliability metrics and would be needed in any online implementation of a commercialized unit commitment package. Of the two formulations, one retains the completely decentralized structure of

the original decentralized unit commitment algorithm proposed here. The other formulation requires certain network sensitivity parameters, which can only be directly calculated with global system information. Following this centralized calculation, the decentralized unit commitment algorithm allows for distributed computation to proceed afterwards. We compare the performances of the two formulations on the basis of computational time and the quality of the solutions found.

The structure of this dissertation is as follows. Chapter 2 reviews some work done in developing a decentralized power systems framework consistent with emerging industry trends and needs. The emerging need for decentralized power system operations is motivated, which leads into the development of decentralized unit commitment in Chapter 3. A formulation of the decentralized unit commitment problem is given along with the application of a solution algorithm. The scalability of the algorithm is demonstrated on a large test system with hundreds of subproblems solved on a computing cluster. The algorithm demonstrated computational speed and objective function optimality gaps well within expected tolerances for the problem size. The value of this computational speed benefit for system operators is discussed and evaluated in Chapter 4. A framework is developed to assess for general cases the value of conducting unit commitment with a faster periodicity to make use of more reliable information close to decision-making deadlines. Chapter 5 further refines the decentralized unit commitment formulation and algorithm with an ability to handle network contingencies, a crucial security and reliability function of realistic unit commitment processes. Finally, Chapter 6 summarizes the novel contributions made through this work and delivers concluding remarks.

CHAPTER II

DECENTRALIZED POWER SYSTEMS ARCHITECTURE

Portions of the following section are quoted or paraphrased from [34], which is available upon request.

2.1 Motivations

The electricity industry's centralized control paradigm has so far been successful in achieving reliable electricity service at reasonable cost. Emerging trends, such as secure integration of unpredictable and variable renewable energy, deployment of smart grid sensing and communication infrastructure, and new consumer objectives result in substantially amplified communication, data management, and computation requirements [19]. As the number of intelligent devices on the grid continues to grow, it becomes increasingly important for those devices to be coordinated at the system level to maintain secure operation.

As an example of the problems that can be caused by lack of coordination, consider electric vehicle (EV) charging. When large numbers of EVs are plugged in at the same time and drawing power from the distribution grid, protective devices may operate incorrectly and transformer life may be reduced. In simulations of uncoordinated EV charging conducted in [15], the authors found that power quality suffered due to significant voltage deviations in high-demand periods. Problems of uncoordinated operation arise not only from EVs but also from photovoltaic (PV) installations [82], demand response [49], and other distributed energy resources (DER) [65].

2.2 Limitations of Centralized Architecture

Traditionally, real-time coordination of multiple devices for system operation has been conducted in a centralized way. Data from measurement units deployed across large systems are aggregated to a single database to assemble a global picture of the system state with which decisions can be made about how to actuate each device to achieve the control objective. This approach provides humans in the loop with clear reporting and high confidence that control decisions are being made correctly. The price of this confidence is seen in the rising costs of communication and data management systems that must scale up with the introductions of devices like DERs and PMUs. Further, the confidence that humans can have in the system state and in the safety of their control actions is eroding as the data becomes more difficult to analyze and decisions must be made more quickly.

The centralized architecture suffers from fundamental scalability limitations due to the following factors:

- Expanding data requirements: The number of monitoring and control devices is increasing by several orders of magnitude over the past, resulting in a control center dilemma between coarsely granular data and a tsunami of information.

- Communication bottlenecks: Centralized control will require increasingly heavy use of communication systems, especially as phasor measurement units are deployed. Expensive, dedicated communication systems will be needed.

- Intractable control and optimization problems: Most current forms of stochastic optimization result in problem sets that are intractable in the time frame required for operation. Further, contingency analysis can become intractable for large systems due to the number of scenarios and sensitivities that must be examined.

- Risks of large-scale renewable energy: A high penetration of renewable energy poses operational challenges and can result in system contingency events.
- Growing complexity of system operations: The number and complexity of reliability and compliance procedures is growing rapidly as the industry integrates renewable energy and addresses concerns such as inter-area oscillations, the effects of demand response, and deployment of energy storage.
- Cyber-security: Centralized control remains a cyber and physical security target. Bulk energy control centers require major infrastructure to be physically protected and usually redundant facilities, hardware, and software infrastructure.
- Data privacy: A centralized framework results in the central organization controlling non-owned assets, requiring significant amounts of data to be transferred from those non-owned control points. The sensitivity of data has been a major concern of ISO members since deregulation began. Data privacy issues have been pointed out in smart grid pilots as well and have resulted in pushback from consumers.

Thus, there is a need for an evolved model for managing the electricity infrastructure that reduces complexity, enables decentralized decision-making, allows for more flexible control, and supports desirable value propositions.

2.3 New Objectives

The need for a new framework is correlated with two changes in industry objectives:

- The traditional objectives of reliability and economic optimization are becoming more difficult to achieve due to infrastructure operating nearer to its limits.
- New objectives of sustainability, national energy security, and support for energy services are growing in importance.

The U.S. Department of Energy identified in [19] seven objectives of the smart grid from a review of stakeholder perspectives. A mapping between these objectives and the five electricity industry objectives mentioned above is provided in Table 2.1. It should be noted that several of the smart grid objectives have impact primarily in the newer industry objectives of sustainability, energy security, and energy services. Realizing these objectives has been recognized by many industry observers as not only an engineering challenge but an architectural challenge [24, 66]. The core elements of the framework proposed by [34] are summarized next.

2.4 Elements of Decentralized Grid Architecture

2.4.1 Element 1 – Distributed Decision-Making

In the centralized architectural framework, there is one system-level objective function corresponding to minimization of total operating cost of a given geographic region. Producers, consumers, and distribution utilities yield control of their assets to the system operator who conducts centralized optimization and control. In the smart grid arena, users may find long-term benefit from installing local generation (e.g. rooftop solar or energy storage) to hedge against real-time prices or pursue sustainability goals. Because of the diversity of objectives that power grid actors have and the common goal of maintaining a reliable system, a control platform should optimize individual behaviors in concert to achieve system-level objectives. Ultimately, the disparate classes of producers and consumers will evolve into economically motivated entities that pursue their own long- and short-term energy objectives. The management architecture must provide the platform to support such distributed decision making.

2.4.2 Element 2 – The Prosumer Abstraction

Clients of the electricity grid can be modeled under a scale-free cyber-physical abstraction. For this abstraction, we introduce the term *prosumer*, meaning an economically

Table 2.1: Smart Grid and New Electricity Industry Objectives

Smart Grid Objectives	Econ.	Reliability	Sust.	Energy Sec.	Serv.
Self-Healing		●		·	
Consumer Empowerment	·		·		●
Power Quality		●			
Tolerant of Attack		●		·	
Variety of Generation Options			●	·	
Maturing Electricity Markets	●				·
Optimize Assets	●				

motivated entity capable of producing, consuming, storing, or transporting electricity. The prosumer is an electric power system with an owner or operator who establishes an energy-related objective function for the system. It is equipped with sensing, communication, and decision-making logic that assists it in pursuing that objective. The prosumer abstraction, originally developed in [35, 36], is illustrated in Fig. 2.1, which shows a generic power system with connections to external supply, local production and storage, and some loads. Electric power systems exhibit similar physics at various scales. They are all electric circuits subject to Ohm's and Kirchhoff's laws. Prosumers can therefore be generalized to power systems of any scale, expressing the interactions among various power systems as interactions among prosumers.

One illustrative example of a prosumer is an electric vehicle with three interaction modes. While disconnected from the grid, it may serve a specific objective determined by the owner, such as to optimize for energy-efficient driving. When connected at home, the vehicle may yield its objective function to the home and instead send a constraint, such as the need to be fully charged by 5:00 AM the next day. Hence, the vehicle becomes an asset of the home prosumer. Similarly, when connected to an office parking lot, the vehicle prosumer may impose charging constraints. The vehicle may also be providing other prosumer services such as deferrable load frequency regulation. Fig. 2.2 summarizes these scenarios.

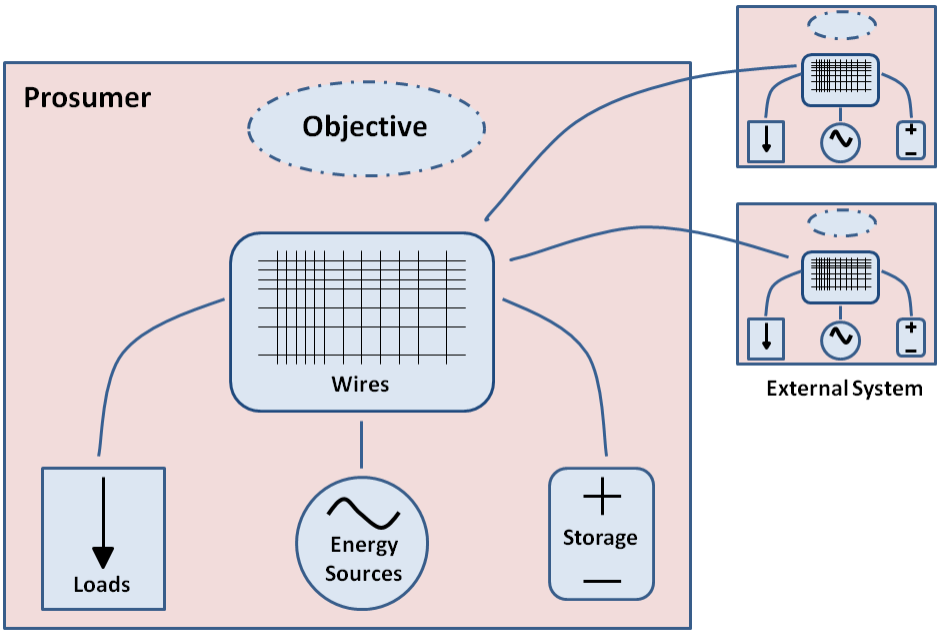

Figure 2.1: Prosumer abstraction to generic elements.

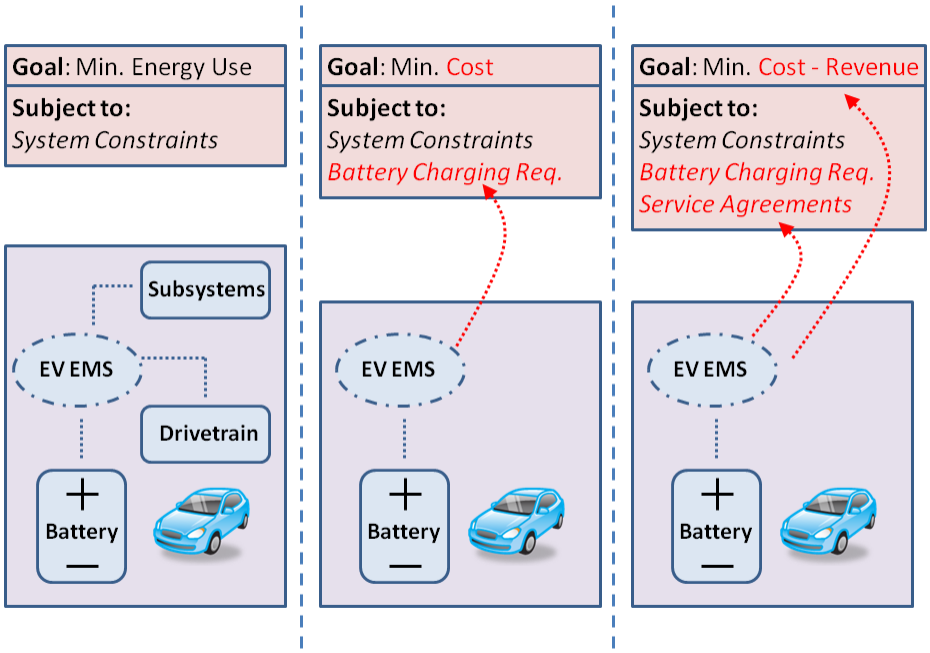

Figure 2.2: Electric vehicle prosumer interactions.

Table 2.2: Power Balancing Time Scales

Time Scale	Driver	Mechanism
24-hr and longer	Seasonal planning Provider decisions	Contracts
30 min – 24 hrs	Human activity Energy storage	Energy scheduling (UC and ED)
30 sec – 30 min	Wind/solar variability	Reserve services
30 ms – 30 sec	Machine dynamics	Frequency Regulation
$<$ 30 ms	Power electronics	Fast Freq. & Voltage Regulation

2.4.3 Element 3 – Coordinated Temporal Scales

The fundamental reliability concern of system operation is making sure power generation is balanced with power demand. In the long term, the power balancing problem is solved by the actors engaging in power exchange contracts. This is true for both large interconnections and for a homeowner who has a contract with the local utility. In the day-ahead time frame, a co-optimized market is used to determine generation, load, and reserve schedules of balancing entities within an ISO. Currently, smaller prosumers do not contribute to power balancing. Dynamic pricing, demand response, distributed storage, and frequency regulation programs by smaller prosumers are clear indicators of a trend towards procurement of services to balance power using distributed resources at a variety of temporal scales. Table 2.2 lists the various time scales and the power balancing methods associated with each.

2.4.4 Element 4 – Distributed Autonomous Control

Recent advances in distributed networked control and multi-agent theory provide a formalism under which the prosumers of the future grid will interact as autonomous agents while observing system-level constraints, which can be enforced in a decentralized manner. Results from [68] demonstrate that prosumers can, in a decentralized fashion, reach a state of agreement on power interchange that respects power balance and minimizes deviation from prosumers' desired energy consumption.

2.4.5 Element 5 – Prosumer Services Cyber-Infrastructure

The prosumer as a generic power system entity with the tasks of generation, consumption, storage, and transportation is simplified by the prosumer fulfilling the fundamental task of balancing power. The examples mentioned above can all be framed as seeking energy balance over a period of time. To external prosumers, imports and exports are abstracted from the specific generation, consumption, and storage capabilities. The concept of virtual power plants [55] already uses an abstraction of generation services, which can be obtained from generation, demand response, discharging storage systems, or a combination.

2.5 Organization into Layers

The core elements described above can be arranged as a layered architecture to allow stakeholders to design, test, and implement decentralized power system control technologies without having to design an entire framework for each effort. A high-level view of the proposed architecture layers is shown in Fig. 2.3.

The device layer corresponds to electrical power devices themselves: generating machines, solar panels, transformers, batteries, and any other electrical supply equipment. It involves core high-voltage transmission technologies, but it also includes low-voltage wiring, light bulbs, and the electrical components of appliances such as motors.

The local control layer includes the hardware and software used for controlling devices (e.g. generator governor or a battery charger). This layer involves the device instrumentation, sensors, actuators, and embedded algorithms for local control and protection devices.

The cyber layer consists of communications, information, and computation infrastructures. It is a platform that supports control logic and economic decision making at the system level.

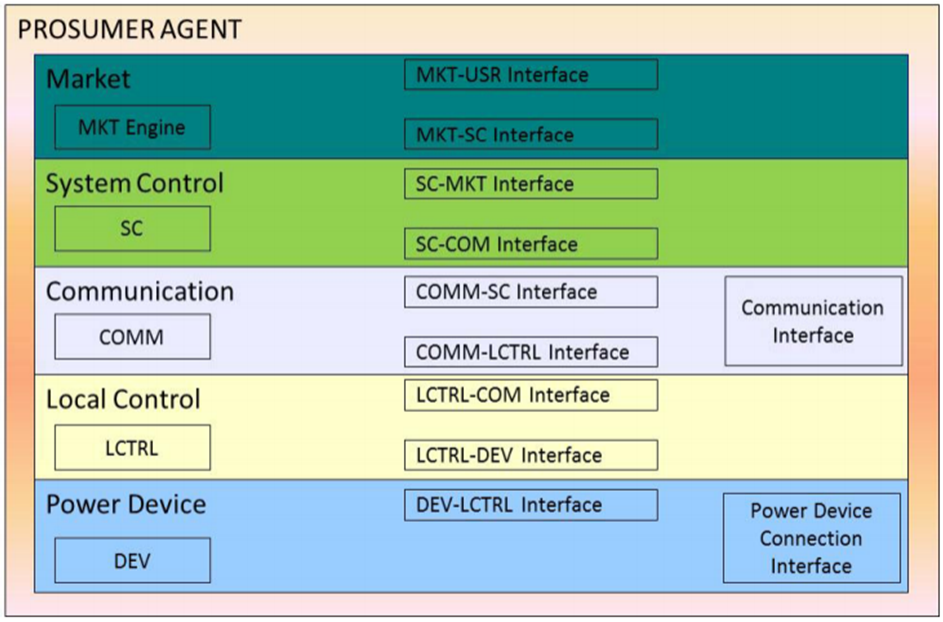

Figure 2.3: Layered architecture of each prosumer.

The system control layer addresses the system security and reliability functions of the grid. It involves two aspects: control internal to the prosumer (monitoring, optimization, and internal network communication) and coordination of prosumer actions with the wider system. The internal system control may include some of the functions currently available in energy management systems, such as state estimation or security assessment. The external function addresses developing decentralized reliability protocols. The UC algorithm described in Chapter 3 would fall under the purview of this layer.

The market layer consists of two aspects as well. The internal aspect addresses the satisfaction functions that realize objectives of economics, sustainability, and support for energy security and energy services. The external aspect addresses interactions concerning service prices and the interpretation of price signals, service transactions, and strategic decision making.

CHAPTER III

DECENTRALIZED UNIT COMMITMENT

3.1 Introduction

The following chapter is based on work performed jointly with Mohammad Javad Feizollahi. This work is described in reference [27], which Mr. Feizollahi and Mr. Costley co-authored, and it is largely reproduced here. The contributions of the overall work are discussed in Sections 3.1.4 and 3.5. Mr. Feizollahi's contributions included the perception variable approach of Section 3.2.8, application of ADMM to handle constraints (32), and the R&F process described in Sections 3.3.3 and 3.3.4. Mr. Costley's contributions included the approach to handle system reserve constraints (17), which is described in Section 3.3.1, and the large-scale software systems to conduct the centralized and decentralized experiments in Section 3.4.

3.1.1 Motivation

Unit commitment, the problem of scheduling electric power plant operations over some time horizon subject to economic concerns, is one of the central operations problems posed by system operators. At its core, the unit commitment (UC) problem is a mixed-integer program (MIP), meaning it is a problem of the form

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{b}_l \leq \mathbf{x} \leq \mathbf{b}_u \\ & x_j \in \mathbb{Z}^+, \forall j \in \{1, 2, \ldots, M\} \end{aligned}$$

where $\mathbf{x} \in \mathbb{R}^{N\times 1}$ with the first M elements integer. In the UC problem specifically, the integer variables are the on/off decisions (represented by values 1 or 0 respectively)

of the generating units. These problems are computationally challenging in general because of the combinatorial solution space of integer variables. The development of high-performance commercial MIP solvers such as Gurobi and IBM CPLEX has made the solution of MIP problems achievable in reasonable computational time. The adoption of branch-and-cut MIP solution methods for solving UC problems is now widespread.

Nevertheless, UC problems continue to increase in difficulty today. This is due in large part to a growing number of combined-cycle natural gas units that can be operated cheaply, meaning optimal schedules involve turning on and off (and switching between various configurations of) these generators at many time steps instead of only operating a few of them at peak load as in the past. This phenomenon is most evident in the market operated by the Midcontinent Independent System Operator (MISO). After the integration of the southern region, the UC engine has begun encountering performance challenges [13]. This can be a significant problem for rolling horizon UC (RHUC) since that is a more time-critical operational procedure than day-ahead UC, and RHUC is responsible for guiding the system when significant deviations from forecast demand or intermittent generation occur. These problems motivate a need for methods scalable to large systems of several thousand buses that can find solutions to UC problems more quickly than the current MIP approach.

Other challenges are emerging besides the size of the system. Growing demand and penetration of wind energy is resulting in congestion along transmission corridors, meaning more binding thermal line constraints must be considered in the UC problem. Additionally, more transmission contingencies must be considered when scheduling generation because of the increase in congestion. Along the boundaries between control areas, the effects of contingencies in one system are not always known in the neighboring system. Therefore, a need is evident for a UC solution method that can include more line constraints than current methods and that can support a more

decentralized information architecture.

As the electricity grid moves toward an architecture like that described in chapter 2, the distribution grid will likely evolve before the transmission grid because of the need on the distribution side to coordinate the operations of many more small discrete devices down to individual residences. However, transmission operations will need to interact with these distribution operations, and decentralized methods for UC may be needed to respond to distribution grid evolution. Such an architecture is proposed in [45].

These factors motivate a UC solution method with the following features:

- Finds good feasible solutions to UC faster than current MIP approach.
- Includes larger proportion of line thermal constraints than is practical today.
- Supports a more decentralized architecture.

The discussion to follow presents the formulation for and discusses the performance of such an algorithm. It is not claimed that this algorithm in its current implementation fully solves the problems just discussed, but because of its fulfillment of the key features listed above, it is promising in its capability to do so. Before developing the formulation, we will summarize the relevant literature, compare this work to the prior research, and propose the novel contributions.

3.1.2 Price-Based Self-Commitment

At the height of the deregulation movement in the United States, a literature around decentralized UC (DUC) sought to investigate the market dynamics of UC conducted by generating companies only (so-called *self-commitment*). In this framework, the independent system operator (ISO) posts hourly energy prices calculated based on the load forecast. Generating companies then conduct UC for their assets as price-takers [78, 70]. Additional iterations may occur to search for an equilibrium where

all the load is served with minimal cost. The existence of such an equilibrium under certain conditions was established in [59, 60] and investigated empirically in [17]. In simulation results reported in [72], the cost of anarchy under self-commitment was found to be roughly 4%. Challenges related to the nonconvexity, oscillation, high cost, and inefficiency of self-commitment were explored in [29, 71, 72, 32].

In this work, we do not study self-commitment as described above. The UC process presented here differs from self-commitment DUC in that it is a direct translation of the traditional centralized UC problem to subproblems corresponding to partitions of the system (which may in turn correspond to generating companies or balancing authorities). In this way, the work in this paper differs in objective and approach from much of the previous work on DUC. Works studying optimal self-scheduling and bidding strategies under some forecast of prices, such as [79, 70, 69, 52, 50, 38], are largely concerned with determining generating company behavior in a framework where some other coordinator (e.g. an ISO) determines prices to induce desirable system-level behavior. As a decomposition of the centralized UC, the work presented here is more analogous to, for example, the reliability UC (RUC) described in [42] conducted following the day-ahead market clearing process. The analyses of [17, 78, 59, 60, 29, 71, 72, 32] study the revenue adequacy of market participants and describe pricing approaches to overcome inefficiencies. In short, these works describe how pricing should occur and market participants should respond in a system without centralized UC, whereas the process described here retains the centralized UC functionality but decentralizes its solution. It retains some of the advantages of self-commitment by ensuring the privacy of commercial data – generating companies need not submit any bids or equipment data to a central agency. Further, it differs from most of the models provided in the works mentioned above by including not only all the traditional generator temporal constraints, but also network flow constraints, which self-commitment cannot easily address. Subproblems can truly be solved in

parallel with minimal requirements for information exchange in each iteration. Although it is certainly possible to implement this in a centralized scheme, it allows for entirely separate entities to coordinate their operations even without a strong centralized computational node.

3.1.3 Decompositions for Unit Commitment

To the extent that DUC is deployed in a single computational environment in a centralized framework, it can be compared to other problem decompositions that leverage parallelizable subproblems. Some of the most important decomposition techniques for UC are Benders [53, 1], dual [28], Lagrangian [54], and Dantzig-Wolfe [30]. These decompositions use a master-slave architecture where subproblems may be solved on separate computational nodes but are coordinated by a master problem, which then requires the results of those subproblems to solve an iteration of its own algorithm.

Other problem decompositions are more similar to that presented here in that they have a decentralized structure without being a form of self-commitment DUC discussed above. Batut and Renaud first applied a regional decomposition approach with duplication of variables to power system problems in [4]. Kim and Baldick [47] similarly used a linearized augmented Lagrangian approach along with the auxiliary problem principle to solve OPF problems in parallel. They then showed in [48] how to extend this formulation to use several different solution algorithms, including ADMM. They note in Section III.B that ADMM is limited in parallel applications because of the interdependency of the two minimization problems. The ADMM method of solving OPF problems in a distributed manner was implemented and tested in [14], but the UC master problem after the Benders decomposition remained centralized. ADMM is also applied to solve security-constrained OPF with AC constraints in [21]. In [62], a Lagrangian method was used to solve the multi-area OPF problem with AC

constraints. The solution computations were largely distributed with a central coordinator needed for some simple calculations in each iteration. An extension of [62] was presented in [33] to solve the decentralized AC power flow using neural networks to solve the nonlinear programming subproblems. Bakirtzis and Biskas [3] proposed a solution to the DC-OPF problem using a similar formulation to [47] around phase angle variable duplication but with a nearly fully decentralized solution approach. Recently, [81] demonstrated a decentralized solution to the security-constrained DC-OPF using a marginal equivalent decomposition that requires exchanging shift factors and binding constraint data in each iteration. This method was proven to converge in finite iterations to the global optimum under some weak assumptions. Finally, [51] provided a DUC framework using a two-level decomposition where regions solve optimization subproblems and communicate to coordinate marginal prices on borders. The formulation uses shift factors, requiring a centralized computation of line-injection sensitivities, and proposes that heuristic methods be used to find feasible solutions to the binary variables.

3.1.4 Contributions

In this paper, ADMM is applied to UC. ADMM is an augmented Lagrangian method similar to dual decomposition but including a quadratic penalty term in the objective function of the subproblems. ADMM as a solution approach has the benefits of augmented Lagrangian techniques, particularly convergence for problems without strict convexity (as we have with the piecewise linear objective functions from generation bids). Furthermore, the approach maintains the separability of the objective function (which motivated the application of decentralized optimization in the first place) in spite of the nonseparable quadratic term. A detailed discussion of ADMM can be found in [6], where it is shown that ADMM converges to a globally optimal solution for convex problems. No such guarantee exists for nonconvex problems such as UC. A

new heuristic process is described to mitigate the nonconvexity presented by binary variables in the DUC problem.

The formulation presented here can be implemented in a peer-to-peer framework that limits information exchange between subproblems, enabling a decentralized structure while preserving the confidentiality of data internal to the regions. In this way, we improve upon the form of ADMM given in [48]. We also address the UC problem, which contains integer variables, as opposed to the OPF problems addressed in [47, 48, 21, 62, 3, 81]. Our formulation differs from [4] in that we solve mixed-integer quadratic programming (MIQP) subproblems instead of finding generation schedules through dynamic programming. It differs from [51] in that we use an augmented Lagrangian formulation and a new heuristic for finding feasible binary solutions. Further, we demonstrate the performance of our algorithm on systems of over 3,000 buses.

The remainder of the chapter is organized as follows. In Section 3.2, we review the notation and problem statement for UC and introduce a formulation suitable for DUC. In Section 3.3, we propose the application of ADMM to network-constrained UC and discuss its challenges. We also provide some heuristics and improvements that address these challenges, resulting in a new algorithm called "release-and-fix". In Section 3.4, we present experimental results for several test systems both in the centralized mode under three models and the decentralized mode partitioned in varying numbers of regions. The main conclusions are summarized in Section 3.5.

3.2 Conventional Unit Commitment Formulation

3.2.1 Introduction

UC is one of the primary functions required by a system operator. In the context of this paper, its objective is to minimize total system costs over the decision variables

of generator active power outputs and on/off status. Constraints mainly include technological aspects. Regulatory constraints include various forms of operating reserve and emissions limits as well as contingency constraints to protect the system against the loss of some generation or transmission elements.

Let the power system network be a connected graph with $\mathcal{N}$ as the set of nodes and $\mathcal{E}$ as the set of edges. In this formulation, nodes correspond to buses and edges to branches (that is, transmission lines or transformers). Buses may have any number of generators and loads connected to them. Suppose the power system is partitioned into n exclusive regions; that is $\mathcal{N} = \bigcup_{\nu \in \mathcal{P}} \mathcal{N}_\nu$ and $\mathcal{N}_\nu \cap \mathcal{N}_{\nu'} = \emptyset$ for all $\nu, \nu' \in \mathcal{P}$, $\nu \neq \nu'$, where $\mathcal{P}$ is the set of regions.

Here we recast the tight MIP formulations for the UC problem presented in [57, 63, 67, 43] with some slight adjustments. These adjustments are as follows. In our formulation:

- The power system is partitioned into regions.
- System reserve requirements are included.
- Both cold and hot startup are considered.
- Quadratic costs for power production are approximated by piecewise linear functions.
- All demand should be served.

In the following, we describe the conventional formulation of the UC problem.

3.2.2 Objective Function

The UC problem seeks to minimize the sum of the total cost of all regions, which is comprised of production, startup, and shutdown costs of all units located in each

region in all times in the study horizon.

$$
\begin{aligned}
C_\nu(\mathbf{x}_\nu) = \sum_{i\in\mathcal{N}_\nu}\sum_{g\in\mathcal{G}_i}\sum_{t\in\mathcal{T}} \big[& C_g^{\mathrm{CS}} v_{gt} + (C_g^{\mathrm{HS}} - C_g^{\mathrm{CS}}) v_{gt}^{\mathrm{HS}} \\
& + C_g^{\mathrm{SD}} w_{gt} + (C_g^{\mathrm{NL}} + C_g^{\mathrm{LV}} \underline{P}_g + C_g^{\mathrm{Q}} \underline{P}_g^2) u_{gt} \\
& + (C_g^{\mathrm{LV}} + C_g^{\mathrm{Q}} \underline{P}_g) p_{gt} + C_g^{\mathrm{Q}} p_{gt}^2 \big] .
\end{aligned} \tag{1}
$$

Cold and hot startup costs are represented using

$$
v_{gt}^{\mathrm{HS}} \le \sum_{\tau=t-T_g^{\mathrm{CS}}-\mathrm{TD}_g}^{t-1} w_{gt}, \quad \forall t \in [T_g^{\mathrm{CS}} + \mathrm{TD}_g + 1, T], \tag{2}
$$

$$
v_{gt}^{\mathrm{HS}} \le v_{gt}, \quad \forall t. \tag{3}
$$

Note that $v_{gt} = 0$ implies $v_{gt}^{\mathrm{HS}} = 0$ by (3), and there will be no associated startup cost in (1). For the case that $v_{gt} = 1$, if no shutdown happened in $[t - T_g^{\mathrm{CS}} - \mathrm{TD}_g, t-1]$, then $v_{gt}^{\mathrm{HS}} = 0$ and the startup cost will be C^{CS}; otherwise $v_{gt}^{\mathrm{HS}} = 1$ and the startup cost will be C^{HS} by (1).

3.2.3 Generator Constraints

Minimum up and down time constraints are represented by

$$
\sum_{\tau=t-\mathrm{TU}_g+1}^{t} v_{g\tau} \le u_{gt} \quad \forall t \in [\mathrm{TU}_g, T], \tag{4}
$$

$$
\sum_{\tau=t-\mathrm{TD}_g+1}^{t} w_{g\tau} \le 1 - u_{gt} \quad \forall t \in [\mathrm{TD}_g, T]. \tag{5}
$$

Startup and shutdown variables are constrained by

$$
u_{gt} - u_{g,t-1} = v_{gt} - w_{gt} \quad \forall t. \tag{6}
$$

Generation limit constraint formulations depend on the generator's minimum up time requirement. If $\mathrm{TU}_g = 1$,

$$
p_{gt} \le (\overline{P}_g - \underline{P}_g) u_{gt} - (\overline{P}_g - \mathrm{SU}_g) v_{gt} \quad \forall t, \tag{7}
$$

$$
p_{gt} \le (\overline{P}_g - \underline{P}_g) u_{gt} - (\overline{P}_g - \mathrm{SD}_g) w_{g,t+1} \quad \forall t. \tag{8}
$$

For generators with $\mathrm{TU}_g \geq 2$,

$$\begin{aligned} p_{gt} \leq (\overline{P}_g - \underline{P}_g)u_{gt} - (\overline{P}_g - \mathrm{SU}_g)v_{gt} \\ - (\overline{P}_g - \mathrm{SD}_g)w_{g,t+1} \quad \forall t. \end{aligned} \tag{9}$$

Note that, as in [57], the variables p_{gt} represent the power generated above the minimum level $\underline{P}_g$.

Ramping limits can be expressed straightforwardly as

$$p_{gt} - p_{g,t-1} \leq \mathrm{RU}_g \quad \forall t, \tag{10}$$

$$-p_{gt} + p_{g,t-1} \leq \mathrm{RD}_g \quad \forall t. \tag{11}$$

The various decision variables included above are constrained as follows:

$$\begin{aligned} p_{gt} \geq 0,\ u_{gt} \in \{0,1\},\ 0 \leq v_{gt} \leq 1 \quad \forall t, \\ 0 \leq w_{gt} \leq 1,\ 0 \leq u_{gt}^{CS} \leq 1 \quad \forall t. \end{aligned} \tag{12}$$

The generator-level constraints for reserve in each hour are as follows:

$$r_{gt}^{\mathrm{SR}} \leq \frac{\mathrm{RU}_g}{6} u_{gt} \quad \forall t, \tag{13}$$

$$r_{gt}^{\mathrm{NSR}} \leq \frac{\mathrm{SU}_g}{6}(1 - u_{gt}) + \frac{\mathrm{RU}_g}{6} u_{gt} \quad \forall t, \tag{14}$$

$$r_{gt}^{\mathrm{OR}} \leq \frac{\mathrm{SU}_g}{2}(1 - u_{gt}) + \frac{\mathrm{RU}_g}{2} u_{gt} \quad \forall t, \tag{15}$$

$$r_{gt}^{\mathrm{OR}} \leq \overline{P}_g - \left(p_{gt} + \underline{P}_g\right) - r_{gt}^{\mathrm{SR}} - r_{gt}^{\mathrm{NSR}} \quad \forall t. \tag{16}$$

The total of each reserve product q in the system must meet the minimal system requirement. For this study, the system requirement was equal to the size of the largest contingency for the 10-minute contingency reserve and the size of the second-largest contingency for the 30-minute operating reserve. Spinning reserve was required to be at least half of the 10-minute reserve. These requirements are based on NERC guidelines discussed in [61]. The system-level constraints are

$$\sum_{\nu \in \mathcal{P}} r_{\nu,t}^{q} \geq \hat{r}_t^q \quad \forall q \in \{\mathrm{SR},\ \mathrm{NSR},\ \mathrm{OR}\},\ \forall t, \tag{17}$$

where $r_{\nu,t}^{q} := \sum_{i \in \mathcal{N}_\nu} \sum_{g \in \mathcal{G}_i} r_{gt}^{q}$.

3.2.4 Network Constraints

Generation must equal demand at each time period, expressed as

$$\sum_{g\in\mathcal{G}} \left[\underline{P}_g u_{gt} + p_{gt}\right] = \sum_{i\in\mathcal{N}} d_{i,t}, \quad \forall t. \tag{18}$$

Note that the constraints (2), (4)-(5), and (7)-(11) should be modified to handle boundary conditions, especially at the beginning of the study horizon. The Appendix of [57] contains the detailed formulations for this.

3.2.5 UC formulation without network constraints

In this formulation, the total cost of the system of regions (1) should be minimized subject to generator constraints (2)-(12) for units in all regions and demand satisfaction relation (18) as follows:

$$\begin{aligned} \min \quad & \sum_{\nu\in\mathcal{P}} C_\nu(\mathbf{x}_\nu) \\ \text{s.t.} \quad & (2)-(16), \quad \forall g\in\mathcal{G}_i, \forall i\in\mathcal{N}_\nu, \forall \nu\in\mathcal{P} \\ & (17), (18) \end{aligned} \tag{19}$$

3.2.6 UC formulation using line sensitivities

In network-constrained UC, thermal limits are considered for branch elements. These are represented as

$$-\overline{F}_{ij} \le F_{ij,t} \le \overline{F}_{ij}, \quad \forall t. \tag{20}$$

Under the DC power flow approximation, the power flow in line ij at time t, which is $F_{ij,t}$, is assumed to be a linear function of net power injections $P^{\text{net}}_{k,t}$ in all buses $k\in\mathcal{N}$; that is,

$$F_{ij,t} = \sum_{k\in\mathcal{N}} \gamma_{ij,k} P^{\text{net}}_{k,t} \tag{21}$$

and

$$P_{k,t}^{\text{net}} := \sum_{g \in \mathcal{G}_k} [\underline{P}_g u_{gt} + p_{gt}] - d_{k,t}. \tag{22}$$

The sensitivities $\gamma_{ij,k}$ are the generation shift factors (GSFs). Denote the matrix of GSFs as $\mathbf{\Gamma}$. We can calculate $\mathbf{\Gamma}$ using the inverse of the admittance matrix, which we call $\mathbf{B}$ under the DC power flow approximation since resistance is ignored. The expression for $\mathbf{\Gamma}$ is

$$\mathbf{\Gamma} = \text{diag}(\mathbf{b})\mathbf{C}^T\mathbf{B}^{-1} \tag{23}$$

where $\mathbf{C}$ is the bus-to-branch incidence matrix of the system and $\mathbf{b}$ is a vector of the branch susceptances.

In network-constrained UC using GSFs, constraints (20) for the monitored lines $ij \in \mathcal{E}' \subset \mathcal{E}$ are added to the model (19). Note that there is no need to actually add the additional variables $F_{ij,t}$ and $P_{i,t}^{\text{net}}$ and constraints (21)-(22) to the optimization model since they can be substituted by the right hand sides of (21) and (22).

$$\begin{aligned} \min \quad & \sum_{\nu \in \mathcal{P}} C_\nu(\mathbf{x}_\nu) \\ \text{s.t.} \quad & (2) - (16), \quad \forall g \in \mathcal{G}_i, \forall i \in \mathcal{N}_\nu, \forall \nu \in \mathcal{P} \\ & (17), (18) \\ & (20), \ (F_{ij,t} \text{ defined as } (21) \text{ and } (22)) \ , \forall ij \in \mathcal{E}' \end{aligned} \tag{24}$$

Because the relation (21) involves a summation across injections from all buses, this constraint is not easily separable or decomposable. We will seek another formulation of branch flows based on the voltages at system buses. Because this formulation adds many variables to the formulation, it is not used in conventional, centralized UC solvers. Because subproblems are much smaller than the global problem, however, formulations including bus voltages are tractable. Next, we develop this decentralized formulation.

3.2.7 Network-constrained UC with voltage phase angles

We can also formulate the DC power flow equations and line limit constraints as

$$F_{ij,t} := B_{ij}(\theta_{i,t} - \theta_{j,t}) \tag{25}$$

and

$$-\overline{F}_{ij} \leq B_{ij}(\theta_{i,t} - \theta_{j,t}) \leq \overline{F}_{ij}, \quad \forall t, \tag{26}$$

respectively, where $\theta_{i,t}$ is the voltage phase angle at bus i and time t. Bus voltage phase angles should satisfy the active power balance constraints

$$P^{\text{net}}_{i,t} = \sum_{j \in \delta_i \bigcup \{i\}} B_{ij}\theta_{j,t}, \quad \forall t. \tag{27}$$

Without loss of generality, bus 1 has been designated the reference bus, giving $\theta_{1,t} = 0$, $\forall t$. Then the network-constrained UC problem can be formulated as

$$\begin{aligned} \min \quad & \sum_{\nu \in \mathcal{P}} C_\nu(\mathbf{x}_\nu) \\ \text{s.t.} \quad & (2) - (16), \quad \forall g \in \mathcal{G}_i, \forall i \in \mathcal{N}_\nu, \forall \nu \in \mathcal{P} \\ & (17) \\ & (26), \forall ij \in \mathcal{E} \\ & (27), \ (P^{\text{net}}_{i,t} \text{ defined as } (22)) \ , \forall i \in \mathcal{N}. \end{aligned} \tag{28}$$

Note that bus balance equations (27) imply the global demand and supply equation (18).

In UC models (19) and (24), the constraints (17), (18), and (20) are globally coupled between different regions, i.e., all of the regions participate in each of these constraints. In model (28), the constraints (26) and (27) are only regionally coupled through boundary lines and buses, respectively. Specifically, they are coupling only between neighboring regions. Next, we propose a reformulation of model (28) that is appropriate to use in our decentralized approach.

3.2.8 Network constrained UC with phase angle perceptions

In model (28), $\theta_{i,t}$ is shared between different regions if i is a boundary bus. For a region ν denote the set of its boundary buses by $\mathcal{N}_\nu^{\mathrm{BB}}$ and the set of boundary buses of the other regions connected to ν by $\mathcal{N}_\nu^{\mathrm{FB}}$. Note that $j \in \mathcal{N}_\nu^{\mathrm{FB}}$ implies $j \in \mathcal{N}_{\nu'}^{\mathrm{BB}}$ for some region ν' a neighbor of ν. See Fig. 3.1 for an illustration of how these three sets relate to the buses of region ν.

To facilitate distributing the UC model among regions, we will assume each region ν connected to boundary bus i has the perception $\tilde{\theta}_{\nu,i,t}$ of the voltage phase angle at bus i and time t. This formulation is similar to that used by Kim and Baldick [47, 48] and the one used by Bakirtzis and Biskas [3], both of which duplicated some variables associated with buses in adjacent regions in order to seek convergence between the regions. However, the formulation presented here differs from the those of [47, 48] by not requiring any dummy buses – all buses belong to some region in the system. Also, this formulation differs from the one in [3] since we use an augmented Lagrangian with different objective terms. Reference [40] similarly uses a variable perception and duplication strategy to formulate an optimal control problem and applies ADMM to solve it in a distributed way.

There are three possibilities for constraints (26) of line ij. If $i, j \in \mathcal{N}_\nu^{\mathrm{IB}}$, then $i, j \notin \mathcal{N}_\nu^{\mathrm{BB}}$ and constraints (26) can be used as-is internal to region ν. For the cases with one internal and one boundary bus as end points of line ij, without loss of generality, let us assume $i \in \mathcal{N}_\nu^{\mathrm{IB}}$ and $j \in \mathcal{N}_\nu^{\mathrm{BB}}$. Then,

$$-\overline{F}_{ij} \leq B_{ij}(\theta_{i,t} - \tilde{\theta}_{\nu,j,t}) \leq \overline{F}_{ij} \quad \forall t. \tag{29}$$

For the cases with one boundary and one foreign bus as end points of line ij, without loss of generality, let us assume $i \in \mathcal{N}_\nu^{\mathrm{BB}}$ and $j \in \mathcal{N}_\nu^{\mathrm{FB}}$. Then,

$$-\overline{F}_{ij} \leq B_{ij}(\tilde{\theta}_{\nu,i,t} - \tilde{\theta}_{\nu,j,t}) \leq \overline{F}_{ij} \quad \forall t. \tag{30}$$

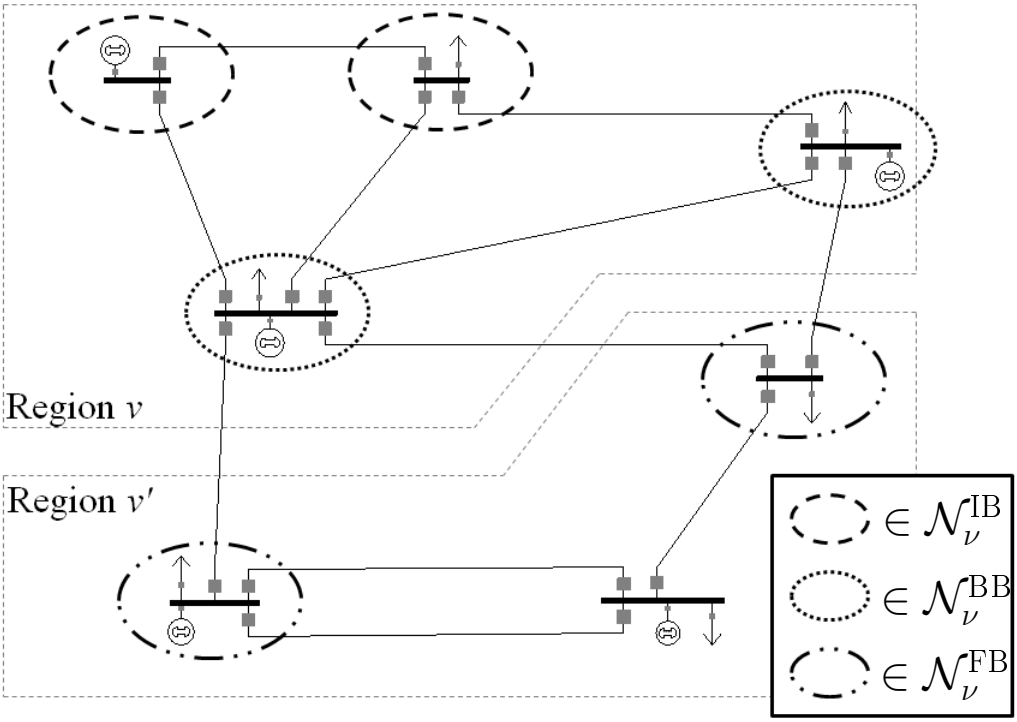

Figure 3.1: Regions ν and ν' outlined with bus sets of region ν.

The power balance constraint (27) for bus $i \in \mathcal{N}_\nu$ can be rewritten as

$$P_{i,t}^{\text{net}} = \sum_{\substack{j \in \delta_i \bigcup \{i\} \\ \text{s.t. } j \in \mathcal{N}_\nu^{\text{IB}}}} B_{ij}\theta_{j,t} + \sum_{\substack{j \in \delta_i \bigcup \{i\} \\ \text{s.t. } j \in \mathcal{N}_\nu^{\text{BB}} \bigcup \mathcal{N}_\nu^{\text{FB}}}} B_{ij}\tilde{\theta}_{\nu,j,t}, \; \forall t. \tag{31}$$

To link the actual phase angles with the perceptions of those phase angles (ensuring agreement on boundary line power flows), we need the additional constraints

$$\begin{aligned} \tilde{\theta}_{\nu,i,t} = \tilde{\theta}_{\nu',i,t}, \quad & \forall \nu, \nu' \in \mathcal{P}, \forall t \in \mathcal{T} \\ \text{s.t.} \quad & i \in \mathcal{N}_\nu^{\text{BB}} \cap \mathcal{N}_{\nu'}^{\text{FB}} \text{ or } i \in \mathcal{N}_\nu^{\text{FB}} \cap \mathcal{N}_{\nu'}^{\text{BB}}. \end{aligned} \tag{32}$$

With these, an equivalent expression of the problem (28) is

$$\begin{aligned} \min \quad & \sum_{\nu \in \mathcal{P}} C_\nu(\mathbf{x}_\nu) \\ \text{s.t.} \quad & \mathbf{x}_\nu \in S_\nu, \forall \nu \in \mathcal{P} \\ & (17), (32) \end{aligned} \tag{33}$$

where

$$S_\nu := \left\{ \mathbf{x}_\nu : \left\{ \begin{array}{l} (2)-(16), \quad \forall g \in \mathcal{G}_i, \forall i \in \mathcal{N}_\nu \\ (31), \; (P_{i,t}^{\text{net}} \text{ defined as } (22)) \; , \forall i \in \mathcal{N}_\nu \\ (26), \forall ij \in \left\{ kl \in \mathcal{E} : k, l \in \mathcal{N}_\nu^{\text{IB}} \right\} \\ (29), \forall ij \in \left\{ kl \in \mathcal{E} : k \in \mathcal{N}_\nu^{\text{IB}}, l \in \mathcal{N}_\nu^{\text{BB}} \right\} \\ (30), \forall ij \in \left\{ kl \in \mathcal{E} : k \in \mathcal{N}_\nu^{\text{BB}}, l \in \mathcal{N}_\nu^{\text{FB}} \right\} \end{array} \right\} \right. \tag{34}$$

Note that in model (33) all of the variable and constraints are local except equations (17) and (32). By relaxing these constraints, augmenting them in the objective function, and using ADMM [6], we can iteratively solve (33) in a decentralized framework.

3.3 Solution Approach

3.3.1 Application of ADMM

First, we will relax constraints (17) and (32). For boundary bus i, let $\bar{\bar{\theta}}_{\nu,i,t}^{m}$ be the average perception by all regions $\nu' \in \Delta_i$ of phase angle $\theta_{i,t}$ at bus i, time t, and

iteration m. That is,

$$\bar{\tilde{\theta}}^m_{\nu,i,t} := \frac{\sum_{\nu' \in \Delta_i} \tilde{\theta}^m_{\nu',i,t}}{|\Delta_i|} \quad \forall i \in \mathcal{N}^{\mathrm{BB}}_\nu \cup \mathcal{N}^{\mathrm{FB}}_\nu, \forall t, \tag{35}$$

where $|.|$ is the cardinality of a set. Moreover, let $\bar{\tilde{r}}^{q,m}_{\nu,t}$ be the total reserve product q provided by region ν in the prior iteration summed with the average system deficiency in product q across all regions. That is,

$$\bar{\tilde{r}}^{q,m}_{\nu,t} := r^{q,m}_{\nu,t} + \frac{\hat{r}^q_t - \sum_{\nu' \in \mathcal{P}} r^{q,m}_{\nu',t}}{|\mathcal{P}|} \quad \forall \nu \in \mathcal{P}, \forall q, \forall t. \tag{36}$$

Note that $r^{q,m}_{\nu,t}$ is not a perception variable like $\tilde{\theta}^m_{\nu,i,t}$ and, as such, need not carry a tilde.

Additional terms can be included for the actual power flows on boundary lines, significantly improving convergence in cases where the entries of the matrix $\mathbf{B}$ vary widely in magnitude. Let

$$\begin{aligned} \bar{\tilde{F}}^m_{\nu,ij,t} &:= \frac{\tilde{F}^m_{\nu,ij,t} - \tilde{F}^m_{\nu',ji,t}}{2} \\ &= \frac{B_{ij}\left(\tilde{\theta}^m_{\nu,i,t} - \tilde{\theta}^m_{\nu,j,t}\right) - B_{ij}\left(\tilde{\theta}^m_{\nu',j,t} - \tilde{\theta}^m_{\nu',i,t}\right)}{2}. \end{aligned} \tag{37}$$

Moreover, let $\lambda^m_{\nu,i,t}$, $\omega^{q,m}_{\nu,t}$ and $\mu^m_{\nu,ij,t}$ be the dual variables corresponding to deviations of $\tilde{\theta}_{\nu,i,t}$ from $\bar{\tilde{\theta}}^m_\nu$, $r^q_{\nu,t}$ from $\bar{\tilde{r}}^{q,m}_{\nu,t}$, and $\tilde{F}_{\nu,ij,t}$ from $\bar{\tilde{F}}^m_{\nu,ij,t}$, respectively. For the sake of simplicity, let $\tilde{\theta}_\nu$, $\bar{\tilde{\theta}}^m_\nu$, r_ν, $\bar{\tilde{r}}^m_\nu$, $\tilde{F}_\nu$, $\bar{\tilde{F}}^m_\nu$, λ^m_ν, ω^m_ν and μ^m_ν be the vectors of $\tilde{\theta}_{\nu,i,t}$, $\bar{\tilde{\theta}}^m_{\nu,i,t}$, $r^{q,m}_{\nu,t}$, $\bar{\tilde{r}}^{q,m}_{\nu,t}$, $\tilde{F}_{\nu,ij,t}$, $\bar{\tilde{F}}^m_{\nu,ij,t}$, $\lambda^m_{\nu,i,t}$, $\omega^{q,m}_{\nu,t}$ and $\mu^m_{\nu,ij,t}$ respectively. Now, let the augmented Lagrangian function of region ν at iteration $m+1$ be

$$\begin{aligned} \mathcal{L}_{\rho,\nu}\left(\mathbf{x}_\nu, \bar{\tilde{\theta}}^m_\nu, \bar{\tilde{F}}^m_\nu, \bar{\tilde{r}}^m_\nu, \lambda^m_\nu, \mu^m_\nu, \omega^m_\nu\right) &= C_\nu(\mathbf{x}_\nu) \\ &+ \lambda^m_\nu\left(\tilde{\theta}_\nu - \bar{\tilde{\theta}}^m_\nu\right) + \mu^m_\nu\left(\tilde{F}_\nu - \bar{\tilde{F}}^m_\nu\right) + \omega^m_\nu\left(r_\nu - \bar{\tilde{r}}^m_\nu\right) \\ &+ \frac{\rho}{2}\left\|\tilde{\theta}_\nu - \bar{\tilde{\theta}}^m_\nu\right\|^2_2 + \frac{\rho}{2}\left\|\tilde{F}_\nu - \bar{\tilde{F}}^m_\nu\right\|^2_2 + \frac{\rho}{2}\|r_\nu - \bar{\tilde{r}}^m_\nu\|^2_2, \end{aligned} \tag{38}$$

where $\rho > 0$ is a given penalty factor. The dual variables are updated as

$$\lambda_\nu^m := \lambda_\nu^{m-1} + \rho\left(\tilde{\theta}_\nu^m - \bar{\tilde{\theta}}_\nu^m\right), \tag{39}$$

$$\mu_\nu^m := \mu_\nu^{m-1} + \rho\left(\tilde{F}_\nu^m - \bar{\tilde{F}}_\nu^m\right), \tag{40}$$

$$\omega_\nu^m := \omega_\nu^{m-1} + \rho\left(r_\nu^m - \bar{\tilde{r}}_\nu^m\right). \tag{41}$$

Note that, from (36) and (41), we see that $\omega_\nu = \omega_{\nu'}$ for all regions ν and ν'.

The primal residuals corresponding to $\tilde{\theta}_{\nu,i,t}^m$, $r_{\nu,t}^{q,m}$ and $\tilde{F}_{\nu,ij,t}^m$ at iteration m are $(\tilde{\theta}_{\nu,i,t}^m - \bar{\tilde{\theta}}_{\nu,i,t}^m)$, $(r_{\nu,t}^{q,m} - \bar{\tilde{r}}_{\nu,t}^{q,m})$ and $(\tilde{F}_{\nu,ij,t}^m - \bar{\tilde{F}}_{\nu,ij,t}^m)$, respectively. The dual residuals corresponding to $\bar{\tilde{\theta}}_{\nu,i,t}^m$, $\bar{\tilde{r}}_{\nu,t}^{q,m}$ and $\bar{\tilde{F}}_{\nu,ij,t}^m$ at iteration m are $\rho(\bar{\tilde{\theta}}_{\nu,i,t}^m - \bar{\tilde{\theta}}_{\nu,i,t}^{m-1})$, $\rho(\bar{\tilde{r}}_{\nu,t}^{q,m} - \bar{\tilde{r}}_{\nu,t}^{q,m-1})$ and $\rho(\bar{\tilde{F}}_{\nu,ij,t}^m - \bar{\tilde{F}}_{\nu,ij,t}^{m-1})$, respectively. Let α^m and β^m be the vectors of all primal and dual residuals, respectively, at iteration m.

Now we can express the final augmented Lagrngian formulation at iteration m similarly to (33), but without (32):

$$\begin{aligned} \min \quad & \sum_{\nu\in\mathcal{P}} \mathcal{L}_{\rho,\nu}\left(\mathbf{x}_\nu, \bar{\tilde{\theta}}_\nu^m, \bar{\tilde{F}}_\nu^m, \bar{\tilde{r}}_\nu^m, \lambda_\nu^m, \mu_\nu^m, \omega_\nu^m\right) \\ \text{s.t.} \quad & \mathbf{x}_\nu \in S_\nu,\ (17),\ \forall \nu \in \mathcal{P} \end{aligned} \tag{42}$$

3.3.2 DUC Algorithm

We propose the basic DUC algorithm as follows.

Step 1. *Initialize primal variables* $\mathbf{x}_\nu^0$ *and dual variables* λ_ν^0 *and* μ_ν^0 *for each region* $\nu \in \mathcal{P}$*; set* $m = 0$.

Step 2. *Each region* ν *sends* $\tilde{\theta}_{\nu,i,t}^m$ *and* $\tilde{\theta}_{\nu,j,t}^m$ *for all time* t *for all* $i \in \mathcal{N}_\nu^{BB} \cap \mathcal{N}_{\nu'}^{FB}$ *and* $j \in \mathcal{N}_\nu^{FB} \cap \mathcal{N}_{\nu'}^{BB}$ *to its neighbor* ν'.

Step 3. *Each region* ν *computes* $\bar{\tilde{\theta}}_{\nu,i,t}^m$*, for all* $i \in \mathcal{N}_\nu^{BB}$ *and time* t *using* (35) *and sends it back to all regions* $\nu' \in \Delta_i \backslash \{\nu\}$*. Moreover, for each boundary line* ij *such that* $i \in \mathcal{N}_\nu^{BB}$ *and* $j \in \mathcal{N}_\nu^{FB}$*, region* ν *updates* $\bar{\tilde{F}}_{\nu,ij,t}^m$ *from* (37).

Step 4. *Each region ν sends $r_{\nu,t}^{q,m}$ for all time t and all products q to a designated region ν^*. Region ν^* calculates $\bar{\bar{r}}_{\nu,t}^{q,m}$, the new reserve targets for each region, by* (36) *and sends the corresponding targets out to each region.*

Step 5. *Each region ν updates its primal and dual residuals. Dual variables are updated using* (39) – (41).

Step 6. *For given primal and dual tolerances $\epsilon^{Pri} > 0$ and $\epsilon^{Dual} > 0$, if $m > 0$, $\|\alpha^m\| \leq \epsilon^{Pri}$ and $\|\beta^m\| \leq \epsilon^{Dual}$* STOP *and output $\mathbf{x}_\nu^m$ as optimal decision for each region ν; otherwise go to Step 6.*

Step 7. *Each region ν updates $\mathbf{x}_\nu^{m+1}$ by solving a self-contained problem of the following form:*

$$\mathbf{x}_\nu^{m+1} = \underset{\mathbf{x}_\nu \in S_\nu}{\operatorname{argmin}} \ \mathcal{L}_{\rho,\nu}\left(\mathbf{x}_\nu, \bar{\bar{\theta}}_\nu^m, \bar{\bar{F}}_\nu^m, \bar{\bar{r}}_\nu^m, \lambda_\nu^m, \mu_\nu^m, \omega_\nu^m\right) \tag{43}$$

Set $m \leftarrow m + 1$ and go to Step 2.

3.3.3 Solution Processes

If the problem (33) has a feasible and bounded optimal solution and the sets S_ν are convex, closed, non-empty sets for all $\nu \in \mathcal{P}$, the approach proposed in Algorithm 1 can solve (33) in a decentralized framework (see [6] for convergence properties of ADMM). However, because the binary variables u_{gt} introduce significant nonconvexity, oscillations in the binary variables may prevent effective searching for the global optimum. Several heuristic modifications to the direct application of ADMM were developed. The most relevant to this discussion are summarized here. Let $\mathbf{u}_\nu$ and $\mathbf{y}_\nu$ denote the subvectors of binary decision variables and continuous decision variables, respectively, in $\mathbf{x}_\nu$.

3.3.3.1 ADMM-CR

ADMM-CR denotes the application of ADMM to the continuous relaxation CR(S_ν) of the set of feasible solutions of each region ν. That is, each of the variables in

$\mathbf{u}_\nu$ is allowed to take any value between 0 and 1. The subproblems then become convex quadratic programs (QPs), for which ADMM is known to demonstrate strong convergence [6]. Solutions from ADMM-CR can provide a good lower bound when the formulation is tight.

3.3.3.2 ADMM-BIN+

ADMM-BIN+ refers to the application of ADMM with MIQP subproblems where the binary variables are required to take 0-1 values. Note that the subproblems always have quadratic objective functions because the augmented Lagrangian (38) is quadratic. The computation time due to the MIQP subproblems can be reduced by observing the transitions in binary variables. If, when solving the MIQP subproblems, some regions do not change the values of their binary variables for some number of consecutive iterations (i.e. some elements of $\mathbf{u}_\nu$ are remaining constant), then we may fix those binary variables temporarily. This is consistent with the empirical observation that only a subset of binary variables are actively being searched at any given stage of the solution process. Furthermore, the penalty factor ρ can be decreased or increased at different points depending on the history of solutions. If new binary solutions are needed, ρ can be decreased to encourage exploration of new binary values. If the ADMM-BIN+ stage has been running for many iterations, ρ may be increased to force settlement on a binary solution.

3.3.3.3 ADMM-BIN–

When a binary-feasible solution is found, the binary variables can be fixed while the solution of continuous variables is refined through further ADMM iterations. This differs from ADMM-BIN+ in that the whole vector $\mathbf{u}_\nu$ is fixed for all ν. Therefore, the only active decision variables are the $\mathbf{y}_\nu$. With sufficiently low residual tolerances, the solution resulting from this algorithm is implementable, unlike ADMM-CR. The strong convergence of ADMM-CR is retained, however, since the subproblems are

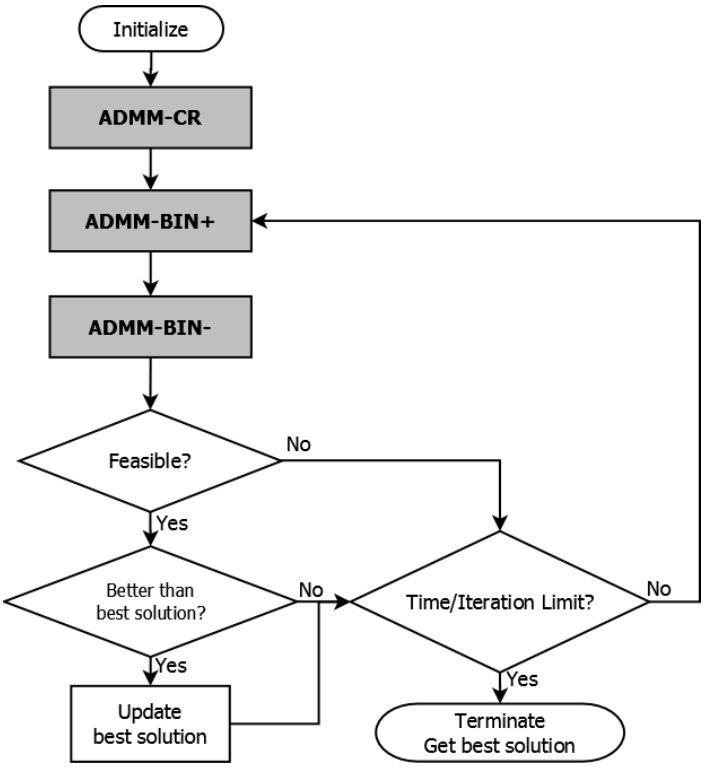

Figure 3.2: Composition of solution processes for DUC.

convex QPs. ADMM-BIN– is analogous to economic dispatch, in which the commitment decisions are generally thought of as fixed while optimization is conducted over the continuous variables.

3.3.3.4 Release-and-Fix process

The flowchart of Fig. 3.2 illustrates the basic process of the Release-and-Fix (R&F) approach. The ADMM-CR stage is used to provide good starting points for many decision variables and the dual variables. The following stage is composed of cycles between the ADMM-BIN+ and ADMM-BIN– processes where binary solutions are explored, discovered, refined, and recorded before searching for more binary solutions.

3.3.4 Improvements

The performance of the algorithms discussed above can be significantly improved by some of the following:

3.3.4.1 Subproblem acceleration

Several tweaks can be made to the optimization software package to further improve solution performance. First, selecting the appropriate root node algorithm for the subproblems can greatly reduce subproblem solution time. When subproblems are smaller, a dual simplex approach is more beneficial than an interior point method, and vice-versa. Furthermore, if a simplex method is used, inheriting the basis from the solution of a previous iteration provides even more speed. In CPLEX, it was found that the full MIP preprocessing was often run for problems in the ADMM-BIN– stage (i.e. all binary variables fixed), even though the problem to solve was effectively a QP. Manually changing the problem type was necessary to leverage this knowledge.

3.3.4.2 Warm start

As will be seen in the experimental results, initializing the subproblems with primal and dual variables from a hypothetical previous day's solution can sometimes reduce solution time. Even if the load forecast is significantly different, resulting in not only a different dispatch but a different commitment, the initial solution can have much smaller primal and dual residuals than starting from the origin.

3.3.4.3 Stopping criteria

Adjusting stopping criteria depending on the problem structure can help balance feasibility with solution speed. For problems with many boundary lines, the magnitudes of the primal and dual residual vectors will naturally be larger than problems with fewer boundary lines. By scaling the total residuals against the totals of the associated variables, more general stopping criteria can be defined.

Table 3.1: Power System Test Cases

Power System	Generator								Total Gens	Capacity (MW)	Buses	Lines
	1	2	3	4	5	6	7	8				
A	19	19	23	25	25	18	8	13	150	30,415	3,012	3,566
B	42	36	12	10	21	14	10	7	152	44,107	3,374	4,068

3.4 Experimental Results

3.4.1 Test Case Details

Two power systems, A and B, were used to conduct experiments to test this algorithm. Some basic information regarding their structures are shown in Table 3.1. Network topologies for the systems A and B are adapted from the IEEE 3,012- and 3,375-bus cases, respectively, available in the MATPOWER software package [83]. Because one of the buses in the 3,375-bus case was not connected to the rest of the network, that bus was removed, leaving 3,374 buses. Moreover, to reduce the number of constraints, parallel lines between buses were replaced by their equivalents.

In the MATPOWER cases, which were originally intended for OPF, most of the data needed for UC such as minimum up and down times (TU and TD), ramp rates up and down (RU and RD), and startup costs (C^{HS} and C^{CS}) are not available. Thus, we replaced those generators with the eight classes of generators used in [10, 63, 57] for UC problems with no network constraints. Data for these eight types of generators are given in Table 3.2. In our experiments, quadratic generation costs have been approximated by piecewise linear costs with five line segments of equal length.

For each power system A and B, there were three test cases. In test cases A_1 and B_1, we considered 24 hourly periods, where the total system demand at each hour is determined as given in Table 3.3. Test cases A_2 and B_2 have also 24 hourly periods, but the total system demand is obtained by shifting the demand values in Table 3.3 earlier by one hour (i.e. hour 1 demand is 65%, hour 2 demand is 62%, and so on). Test cases A_3 and B_3 have 72 hourly periods (three days), where the demand in each

Table 3.2: Generator Data

Gen	Technical Information						Cost Coefficients				
	$\overline{P}$ (MW)	$\underline{P}$ (MW)	TU/TD (h)	RU/RD (MW/h)	T^{Init} (h)	T^{cold} (h)	C^{NL} ($/h)	C^{LV} ($/MWh)	C^{Q} ($/MW^{2}h)	C^{HS} ($)	C^{CS} ($)
1	455	150	8	225	+8	5	1,000	16.19	0.00048	4,500	9,000
2	455	150	8	225	+8	5	970	17.26	0.00031	5,000	10,000
3	130	20	5	50	-5	4	700	16.60	0.00200	550	1,100
4	130	20	5	50	-5	4	680	16.50	0.00211	560	1,120
5	162	25	6	60	-5	4	450	19.70	0.00398	900	1,800
6	80	20	3	60	-3	2	370	22.26	0.00712	170	340
7	85	25	3	60	-3	2	480	27.74	0.00079	260	520
8	55	10	1	135	-1	0	660	25.92	0.00413	30	60

Table 3.3: Total Demand (% of Total Capacity)

Time	1	2	3	4	5	6	7	8	9	10	11	12
Demand	71%	65%	62%	60%	58%	58%	60%	64%	73%	80%	82%	83%
Time	13	14	15	16	17	18	19	20	21	22	23	24
Demand	82%	80%	79%	79%	83%	91%	90%	88%	85%	84%	79%	74%

of the three days is the same as A_1 and B_1, respectively. Distribution of demand among buses followed the proportions of the original load data in the MATPOWER files.

To evaluate the decentralized approach, the systems A and B were partitioned into n regions where $n \in \{20, 30, 40, 50, 60, 70, 75, 80, 90, 100, 120, 150, 200\}$. The partitioning algorithm was a heuristic designed to pursue two objectives. The first was minimizing the total number of boundary lines. Although this problem is NP-hard in general, the heuristic method performed adequately. While constructing each region, the buses with the most connections were annexed first to limit the number of edges on the partition boundary. This objective helped reduce the number of augmented penalty terms in each subproblem. The second objective was to equalize the number of generators in each region, since generators are the main contributor to problem complexity. Subproblems are therefore expected to be of roughly similar

sizes. Cases are easy to imagine in which one region is significantly larger than another in terms of generators or problem size, especially if regions are strictly considered to be utilities, generating companies, or transmission operators. However, in these cases, a decomposition could be performed internal to those large regions to achieve better overall congruity. All test case input files and regional assignments of buses in each partition are online, available at [26].

3.4.2 Implementation Details and Algorithm Parameters

All algorithms were coded in C++ using CPLEX 12.6 through the Concert API. Experiments were conducted on a UNIX cluster with cores rated between 2.0 and 3.0 GHz and addressable memory limited to 4 GB. The cluster machines are primarily a variety of Xeon E5 and X5 models.

Central UC instances were solved using internal CPLEX multi-threading with four cores. Test cases A_1, A_2, B_1, and B_2 were solved a computational time limit of two hours, while test cases A_3 and B_3 had a time limit of ten hours. All cases were solved with a 1% relative optimality gap tolerance. Test cases A_1, A_3, B_1 and B_3 did not have any initial solution, while A_2 and B_2 were warm-started with the optimal solutions of A_1 and B_1, respectively. In both centralized and decentralized methods, the barrier method was used to solve root node problems.

The penalty factor ρ was initialized to a value of 2. Whenever the ADMM-BIN+ stage did not find a new binary solution, ρ was multiplied by 0.95 to encourage more exploration in the space of binary variables. Relative primal and dual residual tolerances ϵ^{Pri} and ϵ^{Dual} and the relative tolerance ϵ^{Obj} between the base objective value and augmented objective value were set to 0.5% for all cases. R&F switched from ADMM-CR to ADMM-BIN+ after reaching all of the ϵ^{Pri}, ϵ^{Dual}, and ϵ^{Obj} tolerances or the iteration limit of 400.

In ADMM-BIN+ where binary variables were released, if there were no changes

Table 3.4: Case Sizes in Centralized Models

	No Network Model			GSF Network Model			Phase Angle Model		
Case	# Bin. Vars.	# Cont. Vars.	# Constr.	# Bin. Vars.	# Cont. Vars.	# Constr.	# Bin. Vars.	# Cont. Vars.	# Constr.
A_1	3,600	28,800	67,375	3,600	29,856	68,431	3,600	101,088	310,807
A_2	3,600	28,800	67,375	3,600	29,856	68,431	3,600	101,088	310,807
A_3	10,800	86,400	207,199	10,800	89,568	210,367	10,800	303264	937495
B_1	3,648	29,184	68,323	3,648	30,120	69,259	3,648	110,160	344,539
B_2	3,648	29,184	68,323	3,648	30,120	69,259	3,648	110,160	344,539
B_3	10,944	87,552	211,123	10,944	90,360	213,931	10,944	330,480	1,039,771

Table 3.5: Centralized Solution Information

	No Network Model			GSF Network Model			Phase Angle Model		
Case	Upper Bnd. ($\$ 10^6$)	Lower Bnd. ($\$ 10^6$)	Gap (%)	Upper Bnd. ($\$ 10^6$)	Lower Bnd. ($\$ 10^6$)	Gap (%)	Upper Bnd. ($\$ 10^6$)	Lower Bnd. ($\$ 10^6$)	Gap (%)
A_1	11.298	11.261	0.33	11.314	11.275	0.34	11.316	11.276	0.35
A_2	11.216	11.155	0.54	11.201	11.168	0.30	11.650	11.169	4.13
A_3	33.577	33.408	0.50	33.605	33.440	0.49	35.771	33.446	6.50
B_1	15.490	15.454	0.24	15.536	15.511	0.16	16.179	15.469	4.39
B_2	15.496	15.461	0.23	15.544	15.516	0.19	16.020	15.489	3.32
B_3	46.524	46.346	0.38	46.627	46.501	0.27	–	46.394	–

in binary variables in the last 15 iterations, or an iteration limit of 50 was reached, or the solution satisfied the ϵ^{Pri}, ϵ^{Dual}, and ϵ^{Obj} tolerances and the binary solution was not previously explored, R&F switched to ADMM-BIN-. In the test cases A_1, A_2, B_1 and B_2, the time limit to solve each subproblem was set to 70, 50, 40, and 30 seconds for the configurations with the number of regions set to 20, 30, 40 and 50, respectively. For the other configurations of these test cases, the time limit was 20 seconds. For A_3 and B_3, the above time limits were multiplied by 2. For each MIQP subproblem, the optimal solution of each iteration was used as a warm start for the next iteration. When the MIQP subproblem was being solved for the first time, since there was no warm start, the time limit was multiplied by 3 to find the first feasible integer solution. The optimality gap tolerance was set to 1% for all MIQP subproblems.

In ADMM-BIN– where binary variables were fixed, if there was no decrease in primal residuals for 50 consecutive iterations, that solution was considered an infeasible solution and discarded. Otherwise, it continued until it satisfied the ϵ^{Pri}, ϵ^{Dual}, and ϵ^{Obj} tolerances or reached 100 iterations. At the end of this phase, if the solution satisfied the tolerances and provided a better objective value than the best one recorded yet, the new best solution was recorded. A total iteration limit of 2,000 was used for the whole algorithm.

The message passing interface (MPI) standard was used to develop the distributed software. Decentralized test cases used $n + 1$ computational nodes where each node used a single core. Each region was assigned to one computational node with the final node being used as a simple coordinator. The coordinator node kept track of the R&F stage, checking stopping criteria such as the ϵ tolerances discussed previously, and recording the feasible binary solutions to keep track of which ones had been explored. Note that most of the information internal to each region does not need to pass through the coordinator.

Table 3.6: Centralized UC Solution Times

	No Net. Model		GSF Model		P. Ang. Model	
Case	t_{RR} (s)	$t_{1\%}$ (s)	t_{RR} (s)	$t_{1\%}$ (s)	t_{RR} (s)	$t_{1\%}$ (s)
A_1	3	625	26	648	32	3,975
A_2	3	870	34	1,566	26	–
A_3	12	6,466	170	26,520	952	–
B_1	2	315	31	822	133	–
B_2	2	76	29	598	108	–
B_3	10	4,161	131	9,068	2,164	–

3.4.3 Numerical Results

To provide a benchmark for the DUC approach, three centralized UC formulations were solved for each test case: one without any network constraints (19), one with network constraints through line sensitivities (24), and one with voltage phase angles (28). The GSF network model included only line constraints which were binding or near binding at the optimal solution of the no-network model. This corresponded to 44 lines for the A cases and 39 lines for the B cases.

Table 3.4 shows the problem size of each test case under each of the models. Table 3.5 shows the best upper and lower bounds discovered with the three models, and Table 3.6 presents the computational time required to obtain those results. Some cases could not be solved to 1% optimality gap under all models. For example, only the root node relaxation (the relaxed linear program giving the initial lower bound) was solved for the three-day case B_3. Some of the cases could be solved within the time limit, but they did not reach 1% optimality gap. In such cases, the time $t_{1\%}$ is blank in Table 3.6.

The centralized phase angle model was able to solve the root node relaxation problem in all cases, but much more time was required than with the other two models due to the much larger problem size. For the A_1 and A_2 cases, the root node relaxation problems required similar computational times, but the times required

to close the optimality gap to 1% diverged. In the A_1 case, such a solution was found in under 4,000 s, but in the A_2 case, no solution under 1% optimality gap was found within the time limit of 7,200 s. Furthermore, integer-feasible solutions with optimality gaps of less than 1% could not be found in any of the other cases in the required time frame (note that this was 36,000 s for the A_3 and B_3 cases).

Note that the lower bounds provide some validation for the models. In the model without network constraints, the lower bounds are the smallest, whereas the lower bounds under the phase angle formulation are the highest. This is exactly as expected since binding line constraints drive up the total system cost. Therefore the GSF model has a higher lower bound than the model without network constraints, and the phase angle model has an even higher lower bound since all line constraints are included.

We now describe the results of the DUC experiments. The results in Tables 3.7 and 3.8 and Figs. 3.3-3.5 are based on averages among five runs for each instance. Tables 3.7 and 3.8 reflect the results of a parallel implementation of R&F to solve cases A_1 and B_1, respectively, with the number of regions varying from 20 to 200. In these tables, "# Region" indicates the number of regions (partitions) for each instance. The columns labeled "# Iter", "t_{best}" and "# Cycle" represent the number of ADMM iterations, clock time in seconds spent to get the best solution, and the number of cycles between ADMM-BIN+ and ADMM-BIN–, respectively. All times reported are averages of actual wall clock times. Five runs were conducted of each problem to normalize for occasional performance differences between machines recruited by the cluster job scheduler, which can alter observed solution times. This is an important consideration since the cluster is not solely dedicated to our problems. The "Cost" and "Gap" columns denote the augmented Lagrangian value in millions of dollars and relative optimality gap (in percentage) of the best solution found in R&F, respectively. The optimality gap is based on the best lower bound (LB) obtained from the centralized

Table 3.7: Case A_1 Decentralized Solutions

# Region	# Iter	t_{best} (s)	# Cycle	Cost ($ 10^6)	Gap (%)
20	949	7,179	16.2	11.304	0.25
30	753	4,256	21.8	11.302	0.23
40	796	3,365	24	11.300	0.21
50	583	556	12.2	11.299	0.20
60	850	905	22.4	11.298	0.20
70	749	520	17.6	11.301	0.22
75	704	449	21	11.297	0.18
80	662	389	10	11.300	0.21
90	831	511	13	11.297	0.19
100	646	364	6.4	11.298	0.20
120	809	483	14.6	11.300	0.21
150	822	414	20.4	11.294	0.16
200	775	441	24	11.300	0.21

UC solutions in Table 3.5 as follows:

$$\text{Gap} = \frac{\text{Cost} - \text{LB}}{\text{Cost}} \times 100. \tag{44}$$

Note that, if the UC problem cannot be solved centrally and there is no lower bound available from the central approach, we can use the optimal value at the end of ADMM-CR as a weaker lower bound.

According to Tables 3.7 and 3.8, for cases A_1 and B_1 with the number of regions $n \geq 50$, an optimality gap of 0.3% was reached in less that 10 minutes. Between 650 and 1,000 iterations and 5 to 25 cycles of ADMM-BIN+ and ADMM-BIN– were needed in R&F. Decompositions into small numbers of regions generally did not perform well. Specifically, the 20 and 30 region instances of case A_1 took more than an hour to reach optimality gaps less than 0.30%, which the 150 region instance achieved in under 7 minutes.

In most of the cases, in the first two cycles of ADMM-BIN+ and ADMM-BIN–, R&F was able to achieve 1% optimality gap. Figs. 3.3-3.5 depict wall clock time up to the point of attaining 1% optimality gap for all cases.

Table 3.8: Case B_1 Decentralized Solutions

# Region	# Iter	t_{best} (s)	# Cycle	Cost ($ 10^6)	Gap (%)
20	983	2,677	20	15.502	0.15
30	860	1,437	16.2	15.511	0.21
40	874	986	22.8	15.514	0.23
50	811	665	20.8	15.505	0.17
60	832	701	23.4	15.510	0.20
70	950	577	21.6	15.510	0.20
75	702	323	9.6	15.495	0.10
80	903	415	25	15.500	0.14
90	927	507	16.8	15.514	0.23
100	912	521	24	15.513	0.22
120	812	402	16.4	15.511	0.21
150	790	382	12.6	15.508	0.18
200	770	378	16.6	15.524	0.29

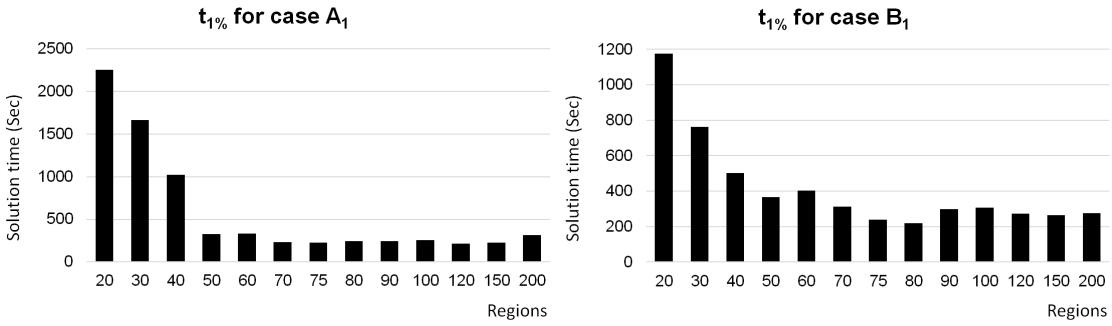

Figure 3.3: Solution time of cases A_1 and B_1.

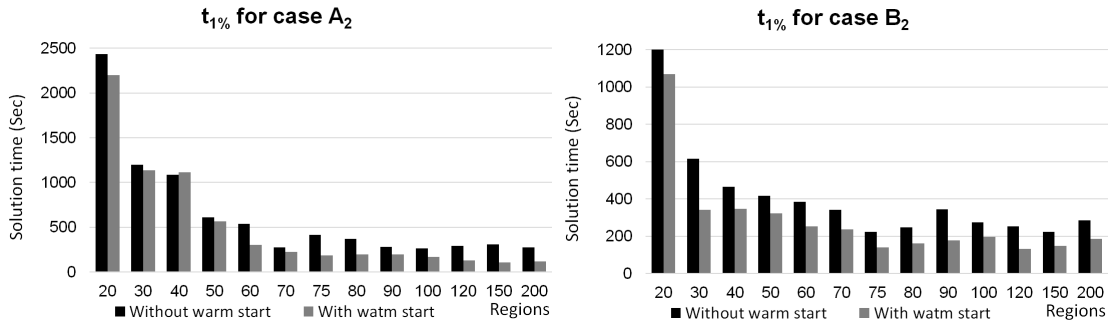

Figure 3.4: Solution time of cases A_2 and B_2 with and without warm start.

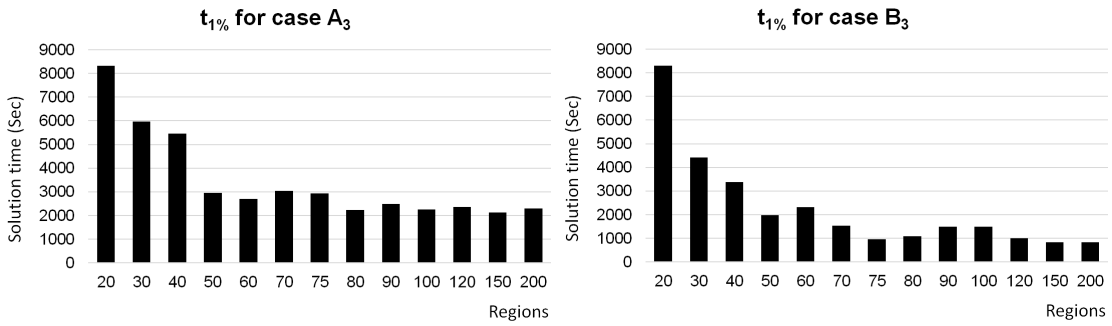

Figure 3.5: Solution time of cases A_3 and B_3.

From Tables 3.7 and 3.8 and Figs. 3.3-3.5, we can conclude that increasing the number of regions to a certain point decreases the solution time, but after such a point, there is no significant decrease (in some cases, it begins to increase). This result is not necessarily surprising. On one hand, ADMM subproblems become easier with increasing numbers of regions, so the computation time for each iteration decreases. On the other hand, the overhead and communication time between regions to update shared values increases by increasing the number of regions. Eventually, this time begins to dominate the gain in computation time of the subproblems in each iteration.

For cases A_2 and B_2, R&F was run both without warm start and with solutions of A_1 and B_1, respectively, as warm-start points. Fig. 3.4 shows a moderate solution time reduction across nearly all the instances of both A_2 and B_2. Because the warm-start solutions were not feasible, R&F had to find totally different commitments for A_2 and B_2.

Fig. 3.5 shows the solution time to get 1% opt. gap for cases A_3 and B_3, which are three times larger than other cases. Solving these cases in a centralized framework was challenging, but R&F found solutions to the corresponding decentralized problem much more quickly. When the number of regions was $n \geq 50$, the 1% optimality gap was reached in less that an hour. For instances of B_3 with greater than 60 regions, this time is less than 30 minutes.

3.5 Conclusions

In this work, the UC problem was formulated in a way suitable for the application of ADMM with some important refinements. The mathematical formulation, addition of new heuristics, and adaptation of parameters based on empirical observations were all studied, tested, and described. The contributions of this work are as follows:

1. Formulation of DUC, including network and temporal constraints, that is scalable up to inter-ISO problems.

2. The R&F approach to deal with binary variables when ADMM is applied to MIQP problems and algorithm parameters providing good performance.

3. Strategies for improving R&F over the basic form.

4. Experimental results demonstrating DUC solution time and solution optimality gaps within reasonable tolerances for large-scale power systems.

The experimental results show that, under a decentralized operations framework, a realistically-sized UC problem can be solved in reasonable time with each region sharing *only phase angle information at boundary buses*. Even in centralized operations, inter-ISO transactions and even near-real-time operations can benefit from a fast, scalable DUC methodology. In fact, the data-sharing requirements of such an application align well with the R&F procedure since no market participant cost data is directly exchanged. Two large ISOs would best leverage the large-scale test results by dividing their interior problems into many regions and then exchanging data on the bordering regions. Because of the poor behavior observed when a small number of large regions are used (see the 20-region test cases), formulating the problem as two ISO-sized subproblems would likely be an inefficient strategy.

Besides the enhancements to data privacy and multi-area coordination offered by R&F to DUC, the computational speed benefit can be exploited by a single ISO with

no change in market or system control architecture. The potential speed increase is enough to open many possibilities, such as:

- Solving MIQP problems to smaller optimality gaps.
- Conducting multiple studies for various scenarios (e.g. under wind power uncertainty).
- Exploring stochastic or robust UC approaches.
- Extending the time horizon or granularity of UC studies.

To achieve a full security-constrained UC in a decentralized setting, some additional steps are needed. A method for conducting contingency analysis without global visibility does not yet exist but is needed even for today's operations. It remains to be determined whether a global contingency analysis can be conducted among decentralized regions that captures full $N-1$ system security. Joint security assessments between neighboring systems are an important topic of ongoing research. Further, the performance of R&F considering AC system constraints should be evaluated. It is hoped that the small subproblems attainable with the DUC formulation might enable reformulation with nonlinear voltage constraints or fast cycling between further subproblems.

It is noted that the applications of a decentralized MIP or MIQP solution methodology extend beyond UC and even power systems. Such an approach might be used to coordinate energy scheduling among a campus of buildings where the number of integer variables, representing states of load operation, may be larger than can practically be solved by other methods. It might also be used to coordinate energy scheduling among a neighborhood of homes aggregated as a demand response unit without concentrating or communicating any data about actual appliance status within the homes. The future electricity grid will require decentralized operations and control

architectures for many reasons outlined in Section 3.1, and the R&F algorithm is designed with such requirements in mind. Given the experimental results obtained so far, it appears that R&F may be able to solve large-scale energy scheduling problems with a highly distributed structure.

CHAPTER IV

ROLLING HORIZON UNIT COMMITMENT

4.1 Introduction

Transmission systems always include a 24-hour UC as part of their operational profile to clear the day-ahead electricity market or, for vertically integrated markets, to synchronize transactions in the day-ahead markets with neighboring systems. However, a 24-hour commitment schedule determined several hours in advance of the operating day is not adequate to operate a system reliably and securely. Imperfect forecasting of load and often intermittent, renewable generation as well as unplanned outages of equipment requires that changes to the UC schedule be made on a consistent basis. If these changes are made with an entirely ad-hoc approach, it is possible to lead the system into insecure states without adequate generation reserves or headroom. At the very least, there is no way to know whether the system is being operated at minimum cost or how close to the minimum cost the system is.

For these reasons, system operators use a so-called "receding-" or "rolling-" horizon approach to operations scheduling where successive problems are solved over and over again as time advances with a constantly progressing look-ahead horizon. Typically, only a subset of the decisions made in each scheduling problem are taken to be fixed, while the others act as "suggestions" for the future. Having a future horizon included in the problem even without any binding decisions in that time frame helps keep the system cost lower over the entire time series of operation and lessens the likelihood of bringing the system to an insecure state. For example, the Midcontinent Independent System Operator (MISO) uses a day-ahead forward reliability assessment commitment (FRAC) and a look-ahead commitment (LAC) before conducting

security-constrained economic dispatch (ED) [12]. The FRAC is conducted from the current time to the end of the operating day while the LAC runs every 15 minutes with a constant rolling horizon. PJM [64] uses so-called "incremental" commitment and decommitment alongside its 5-minute, time-coupled security-constrained ED.

The rolling-horizon approach allows system operators to make use of data that is revealed as time progresses. Especially important for modern transmission grid operations are forecasts for wind generation, which are known to be much more accurate an hour or two ahead of time than a day ahead [39, 8]. If UC could be calculated instantly and generators could receive and respond to actions without delay, the best strategy for system operators would be to wait until just before commitment decisions are needed and decide using the most reliable information. Of course, they are constrained by the calculation time of the problem and by the need to notify generator crews in advance of the schedule.

In this chapter, we will focus on evaluating the benefit that a system operator could glean from being able to calculate UC solutions faster than is done today, demonstrating the value of the computational speed results of Chapter 3. To that end, a platform for simulating operational behavior more realistic than a single UC was developed to simulate the interactions between 24-hour UC, rolling-horizon UC (RHUC), and ED and the improving reliability of forecasting as the prediction time arrives.

There is little prior academic literature on RHUC frameworks or formulations. Tuohy [75] discussed the benefits of the rolling-horizon approach and provided a comparison of the rolling-horizon approach to the static approach both when forecasting was perfect and when it was imperfect. However, even the RHUC in this case was conducted only with hourly granularity, and no explicit formulations are given. Similarly, Constantinescu [16] conducted hourly RHUC, although the time frame was a shrinking interval between the current time and the end of the current operating

day instead of a constant 6 hours ahead of the current time. Feinberg [25] provided some formulations for the RHUC, but only for reserve. Further, RHUC is still only conducted on an hourly basis.

Here, we contribute much more flexible formulations of RHUC and ED by allowing RHUC to be conducted with sub-hour granularity to compare periodicity differences of, in the cases investigated here, 10 minutes. Thus, RHUC could be conducted on 30-minute or 20-minute periods instead of strictly on multiples of 1 hour. Further, we model transitional periods of generator startups and shutdowns in both the RHUC and ED. The framework presented here is therefore capable of a wider range of simulations than those existing and conform more closely to the models of RHUC used in industry.

The remainder of the chapter proceeds as follows. Section 4.2 describes in greater detail the structure of the simulations and the inputs and outputs of each module. Section 4.3 describes the formulation of the RHUC problem, highlighting some important changes in modeling that must be made to couple the 24-hour commitment schedule and retain feasibility of the solution along the operational horizon. Section 4.4 similarly describes the formulation of the ED problem. Section 4.5 describes the test cases and sources of input data, including strategies for generating unavailable data. Finally Sections 4.6 and 4.7 discuss results and conclusions.

4.2 Simulation Structure

Figures 4.1 and 4.2 show the structures of the rolling horizon simulations with 30-minute RHUC and 20-minute RHUC, respectively. RHUC may be done on arbitrary time scales but is generally conducted on a period of less than an hour to take advantage of the sub-hour startup times of natural gas generators. The blocks labeled "UC" and "ED" represent calculations of UC problem and ED problems, respectively, and the length of the blocks roughly corresponds to their calculation time. The solid

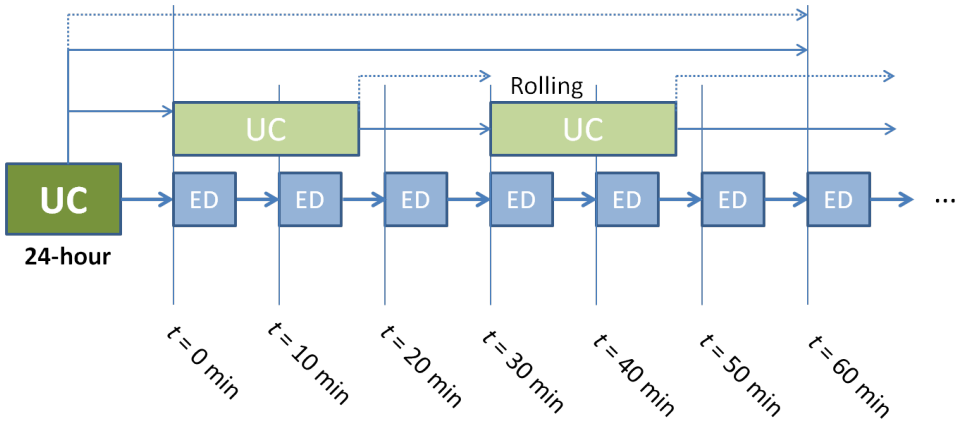

Figure 4.1: Structure of rolling-horizon simulation with 30-minute RHUC.

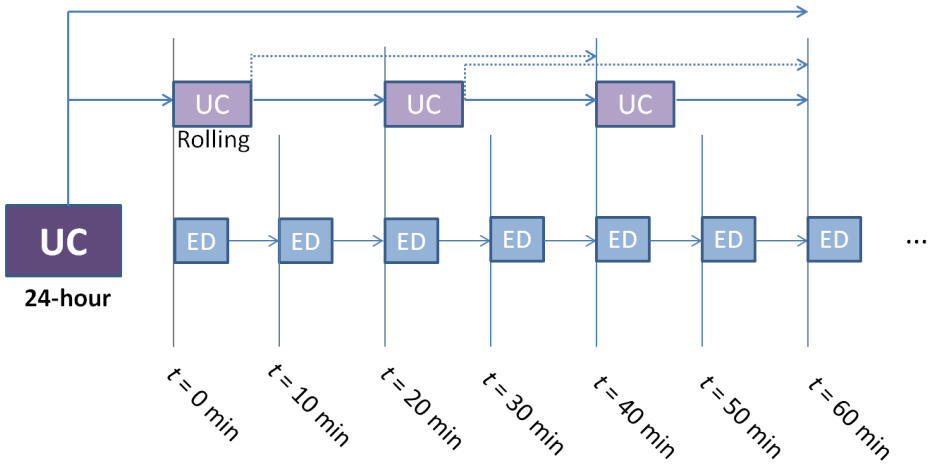

Figure 4.2: Structure of rolling-horizon simulation with 20-minute RHUC.

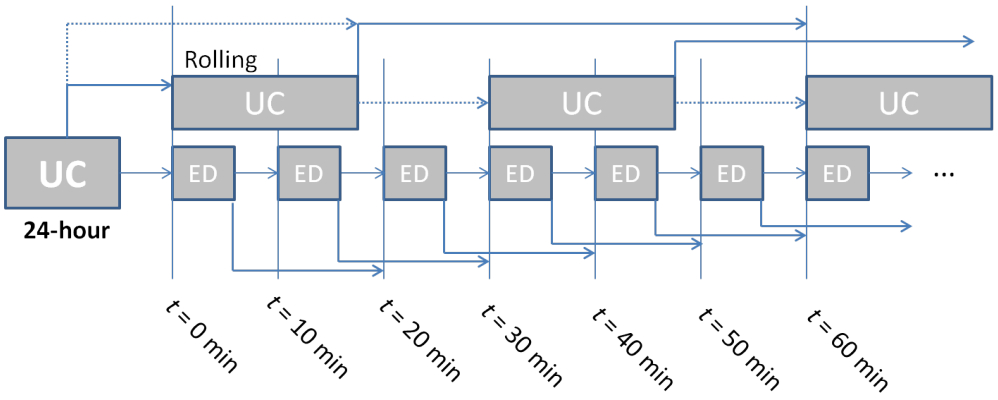

Figure 4.3: Structure of rolling-horizon simulation with 20-minute RHUC.

arrows indicate fixed decisions determined by the UC or ED algorithm at the arrow's source to be applied at the sink. For example, the ED calculated between 0 and 10 minutes creates the fixed power output decisions at $t = 10$ min. Dotted arrows indicate that the output of the UC algorithm at the source has provided a non-binding suggestion for a good schedule of operations at some future time step. The suggested good schedule may be modified by subsequent UC or ED algorithms upon the acquisition of new information, such as an updated load or wind generation forecast. Note that in either of these frameworks, calculations of ED begin at 0 minutes with the result of the solution applied at 10 minutes, with the calculation beginning at 10 minutes eventually providing control action at 20 minutes, and so on.

In most actual operations, the structure includes an added layer of complexity as shown in Figure 4.3, wherein the ED solution calculated between 0 and 10 minutes is applied at 20 minutes. Similarly, RHUC solutions are applied not at the time step immediately following the end of the RHUC calculation, but at the next time step. This affords extra time for the system operator to notify generating companies or plant crews about orders. In this work, we assume the simpler framework of Figures 4.1 and 4.2.

Regardless of the period of the RHUC and ED, a classical 24-hour UC always precedes the operating day. The 24-hour UC will determine the initial operating

point for the operating day as well as the first hour, giving the first RHUC problems start and end points to initialize their set points. Now consider the case in Figure 4.1. Once the operating day begins, RHUC with a 6-hour horizon begins executing at 0 minutes to finish and apply output control actions at 30 minutes. Once those control actions are applied, another UC begins calculation at 30 minutes using the updated load and intermittent generation forecasts available at $t = 30$ min for the new RHUC horizon (30 minutes to 6.5 hours). In between the RHUC calculations, successive ED problems adjust the power outputs of committed generators, the deployment of reserve scheduled by UC, and curtailable loads to react to even newer information and made adjustments necessary to account for generators in their startup or shutdown phases. The inputs and outputs of the various stages are as follows:

- 24-hour UC
  - *Input*: System topology data (e.g. branch impedances, generator costs and constraints, demand data).
  - *Output*: Binding and non-binding UC decisions (UC/UC transfer).
  - *Output*: UC decisions on first UC interval (UC/ED transfer).
- RHUC
  - *Input*: System topology data.
  - *Input*: Updated UC/UC transfer data.
  - *Input*: Updated load/wind forecasts for next 6 hours.
  - *Output*: Updated UC/UC transfer data.
  - *Output*: Updated UC/ED transfer data.
- ED
  - *Input*: System topology data.

- *Input*: Updated UC/ED transfer data.
- *Input*: Updated load/wind forecasts for next 10 minutes.
- *Input*: Updated generation outputs from previous ED.
- *Output*: Control setpoints (e.g. generator power outputs).

Both the RHUC and ED formulations to follow will focus on the centralized problem. Network constraints can be decentralized the same way as in Section 3.2.8.

4.3 Rolling Horizon Unit Commitment Formulation

4.3.1 Preliminaries

Let $\mathcal{T}_h$ denote the time periods that the rolling horizon UC (RHUC) will solve for. Let t_0 represent the time the RHUC optimization procedure starts; that is, the time at which the input information to the UC becomes available. Let $t_1, t_2, ..., t_h$ represent the times in $\mathcal{T}_h$.

For clarity of the formulations here, we assume that all the time intervals are equal to t_{int}; that is,

$$t_{\text{int}} := t_2 - t_1 = t_3 - t_2 = ... = t_h - t_{h-1}.$$

The more general case in which this assumption does not apply only requires that constraints associated with particular time intervals, such as ramping constraints, use the appropriate interval length and that cumulative time calculations be modified in the minimum up and down time constraints. Therefore, although modified constraints would be substituted for each time period, the total number of constraints would not be changed.

Some of the commitment decisions u_{gt} will be fixed from previous UC solutions. If the ordered pair (g, t) corresponds to a fixed commitment decision of generator g at time t, then we say $(g, t) \in \mathcal{D}$ and $u_{gt} = u^*_{gt}$.

Other than slow-start or slow-stop units in transition, we do not consider the power outputs of fixed-commitment generators to also be fixed. That is, p_{gt} is always a decision variable regardless of whether $(g,t) \in \mathcal{D}$ or $(g,t) \notin \mathcal{D}$.

The RHUC will fix commitment decisions for all g for t_1. This is implicit since the commitment at time t_0 would be the last operational procedure to decide on commitment at t_1. Regarding minimum up and down time constraints, we need not explicitly fix the commitment decisions after obtaining the solution. If the parameter T_g^{Init} is updated for each new RHUC based on the behavior of g in the previous interval, then the generators with minimum up and down times will have fixed commitment decisions from the formulation logic already in place.

We distinguish between "slow-start" units and "fast-start" units. To be considered fast-start, a generator must have both $T_g^{\text{CS}} < 1$ and $\text{TU}_g \leq 1$. Similarly, to be considered fast-stop, a generator must have both $T_g^{\text{CS}} < 1$ and $\text{TD}_g \leq 1$. Specifically, we may say that the time required for a fast-start generator to reach the startup capacity SU_g is t_{FS}, where we will generally assume that $t_{\text{FS}} \geq 10$ minutes since that is the maximum time allowed for a generator to deploy its contingency (i.e. spinning or nonspinning) reserve in most systems. Recall that we define the startup and shutdown capabilities as

$$\text{SU}_g := \max\{\underline{P}_g, \text{RU}_g\},$$
$$\text{SD}_g := \max\{\underline{P}_g, \text{RD}_g\}.$$

4.3.2 Objective Function

4.3.2.1 Generator Costs

The generator cost is much the same as the 24-hour UC, except the variable costs are scaled by the time interval t_{int}.

$$\begin{aligned} C_g(\mathbf{x}_g) = \sum_{t_i \in \mathcal{T}_h} \big[& C_g^{\text{CS}} v_{gt_i} + (C_g^{\text{HS}} - C_g^{\text{CS}}) v_{gt_i}^{\text{HS}} + C_g^{\text{SD}} w_{gt_i} + C_g^{\text{T}}(\mathbf{x}_g, t_i) \\ & + t_{\text{int}}(C_g^{\text{NL}} + C_g^{\text{LV}} \underline{P}_g + C_g^{\text{Q}} \underline{P}_g^2) u_{gt_i} \\ & + t_{\text{int}}(C_g^{\text{LV}} + 2C_g^{\text{Q}} \underline{P}_g) p_{gt_i} + t_{\text{int}} C_g^{\text{Q}} p_{gt_i}^2 \big] . \end{aligned} \tag{45}$$

Because $\mathcal{T}_h$ includes time steps during which some generators will be in the process of starting up or shutting down, some units may be outputting power during their transitional period in a period before the startup indicator variable is asserted or after the shutdown indicator variable is asserted. If the full startup or shutdown cost is assessed just on the period when the indicator variables are equal to 1, then the commitment solution will tend to incorrectly favor keeping units in transition to provide free energy. We require transitional startup and shutdown variables to represent these processes for this and other reasons, and they are assessed a corresponding cost represented as

$$C_g^{\text{T}}(\mathbf{x}_g, t_i) := (C_g^{\text{LV}} \underline{P}_g + C_g^{\text{Q}} \underline{P}_g^2)\,(\tilde{v}_{gt_i} + \tilde{w}_{gt_i}) . \tag{46}$$

The transitional startup and shutdown variables $\tilde{v}$ and $\tilde{w}$ are constrained by (55) and (56).

4.3.2.2 Load Costs

We include a cost for load curtailment. This term is simply represented as

$$C_k(\mathbf{x}_k, t_i) := t_{\text{int}} \omega_d (d_{kt_i}^{\max} - d_{kt_i}^{\text{C}}). \tag{47}$$

The cost term ω_d should be much larger than the costs of deploying reserve. We scale the cost term by t_{int} to compare it to the similar term in ED.

4.3.2.3 Penalty Costs

The next part of the objective function is a penalty term for deviating from the schedule prescribed by the 24-hour UC solution. This is necessary to preserve optimality since the RHUC does not consider the entire 24-hour period and therefore may lead the system along different operating trajectories than specified in the day-ahead UC. Given perfect load and renewable generation forecasts, we wish to follow the day-ahead UC as closely as possible since that is the optimal schedule. Penalty terms serve this purpose of keeping the generation schedule close to the optimal inherited schedule. We do not impose any penalty on deviating from decisions in overlapping time steps of previous RHUC schedules.

We penalize only the on/off commitment decisions and not the power output decisions. Rather, the power output decisions as determined by the RHUC are penalized in the ED as explained later. We do not penalize the commitment decisions in $\mathcal{T}_h \setminus \mathcal{T}$ between the 24-hour decision points because, if the generator is able to cycle in less than an hour, it will pay the startup cost as a penalty. We can represent a deviation from the commitment decision as $(u - \hat{u})^2$. Because each of u and $\hat{u}$ must be integer, we can simplify it to

$$(u - \hat{u})^2 = u^2 + \hat{u}^2 - 2u\hat{u} = u + \hat{u} - 2u\hat{u} = \hat{u} + u(1 - 2\hat{u})$$

since the variables u and $\hat{u}$ are binary. Therefore the penalty cost is

$$\Omega_g(\mathbf{u}_g, t_i) = \omega_{\mathrm{U}} \left(\hat{u}_{gt_i} + u_{gt_i}(1 - 2\hat{u}_{gt_i})\right). \tag{48}$$

We include penalty costs that encourage the RHUC to procure the appropriate amount of reserve from the generators. We use $r^q_{t_i}$ to denote the sum of reserve product q provided by each generator in the system at time t_i:

$$r^q_{t_i} = \sum_{g \in \mathcal{G}} r^q_{gt_i}. \tag{49}$$

We scale the terms by t_{int} to compare them to the similar terms in the ED formulation. The objective function includes linear penalty terms on the reserve products with corresponding lower bounds to ensure requirements are met at the RHUC stage (the reserve products may be deployed by the ED as will be seen later). The objective function terms are therefore

$$C_q(\mathbf{x}, t_i) = t_{\text{int}} \omega_q (r^q_{t_i} - \hat{r}^q_{t_i}), \forall q \in \{\text{SR, NSR}\}, \tag{50}$$

and the corresponding constraints would be

$$r^q_{t_i} \geq \hat{r}^q_{t_i}, \forall q \in \{\text{SR, NSR}\}. \tag{51}$$

4.3.2.4 Total Objective

The total objective function for the RHUC problem is then the following quadratic cost:

$$C(\mathbf{x}) = \sum_{t_i \in \mathcal{T}} \left[\sum_{g \in \mathcal{G}} \left[C_g(\mathbf{x}_g, t_i) + \sum_{\substack{t_i \in \mathcal{T} \\ (g,t_i) \notin \mathcal{D}}} \Omega_g(\mathbf{u}_g) \right] + \sum_{k \in \mathcal{L}} C_k(\mathbf{x}_k) + \sum_q C_q(\mathbf{x}, t_i) \right]. \tag{52}$$

4.3.3 Constraints

4.3.3.1 Startup/Shutdown Constraints

Cold and hot startup indicator variables are represented using

$$v^{\text{HS}}_{gt_t} \leq \sum_{\tau = t_i - T^{\text{CS}}_g - \text{TD}_g}^{t_i - 1} w_{gt_i}, \quad \forall t_i \in [T^{\text{CS}}_g + \text{TD}_g + t_1, t_h] \cap \mathcal{T}_h, \tag{53}$$

$$v^{\text{HS}}_{gt_i} \leq v_{gt_i}, \quad \forall t_i \in \mathcal{T}_h. \tag{54}$$

Startup and shutdown variables are constrained by (6) as in the 24-hour UC.

Consider now the case where $t_i \in \mathcal{T}_h$ and $t_- < t_i < t_+$ for some $(g, t_-), (g, t_+) \in \mathcal{D}$. This may happen if, for example, if the decision (g, t_+) is fixed due to scheduled maintenance of the generator g at t_+. If $t_+ - t_- > 1$, then we can leave u_{gt_i} free and

wait until later to fix the decision. Therefore assume $t_+ - t_- \leq 1$. We then have two simple cases:

$$\begin{cases} u_{gt_i} = u^*_{gt_i} = 1 & \text{if } u^*_{gt_-} = u^*_{gt_+} = 1 \\ u_{gt_i} = u^*_{gt_i} = 0 & \text{if } u^*_{gt_-} = u^*_{gt_+} = 0. \end{cases}$$

Assume we have $t_{\text{int}} \geq t_{\text{FS}}$. Then we can write several cases to handle startup and shutdown intervals:

$$\begin{cases} u_{gt_i} = u^*_{gt_i} = 0 & \text{if } u^*_{gt_+} - u^*_{gt_-} = 1 \text{ and } g \in \mathcal{G}^{\text{SS}} \\ u_{gt_i} \text{ free} & \text{if } u^*_{gt_+} - u^*_{gt_-} = 1 \text{ and } g \in \mathcal{G}^{\text{FS}} \\ u_{gt_i} = u^*_{gt_i} = 0 & \text{if } u^*_{gt_-} - u^*_{gt_+} = 1 \text{ and } g \in \mathcal{G}^{\text{SSh}} \\ u_{gt_i} \text{ free} & \text{if } u^*_{gt_-} - u^*_{gt_+} = 1 \text{ and } g \in \mathcal{G}^{\text{FSh}} \end{cases}$$

Fast-start and fast-stop units will be considered fully operational or offline, respectively, within t_{int}, so no special cases are needed for them. For the first and third cases listed above corresponding to slow-start and slow-stop units, the regular generator model will be suspended, and the power output decision $\bar{p}_{gt_i}$ will be fixed to $\text{SU}_g(t_{\text{int}} - t_-)$ or $\text{SD}_g(t_+ - t_{\text{int}})$ depending on whether the unit is starting up or shutting down. In such cases, we will also restrict the reserve to be zero.

We have "transitional" startup variables to represent when slow-start units are in the process of starting up and shutting down. These variables help in establishing the power outputs and costs of the power outputs as the unit ramps up or ramps down during startup and shutdown. We make use of the term

$$n_{\text{SS}} := \left\lceil \frac{t_{\text{SS}}}{t_{\text{int}}} \right\rceil.$$

Transitional startup variables are constrained by

$$\tilde{v}_{gt_i} = \sum_{j=1}^{n_{\text{SS}}-1} \frac{n_{\text{SS}} - j}{n_{\text{SS}}} v_{gt_{i+j}}, \quad \forall i \in \{1, 2, ..., h - n_{\text{SS}} + 1\} \tag{55}$$

with transitional shutdown variables similarly constrained by

$$\tilde{w}_{gt_i} = \sum_{j=0}^{n_{\mathrm{SS}}-2} \frac{n_{\mathrm{SS}} - j - 1}{n_{\mathrm{SS}}} w_{gt_{i-j}}, \quad \forall i \in \{n_{\mathrm{SS}} - 1, n_{\mathrm{SS}}, ..., h\}. \tag{56}$$

4.3.3.2 Minimum Up/Down Times

Minimum up and down time constraints are represented by

$$\sum_{\tau=t_i-\mathrm{TU}_g}^{t_i} v_{g\tau} \leq u_{gt_i} \quad \forall t_i \in [\mathrm{TU}_g + t_1, t_h] \cap \mathcal{T}_h, \tag{57}$$

$$\sum_{\tau=t_i-\mathrm{TD}_g}^{t_i} w_{g\tau} \leq 1 - u_{gt_i} \quad \forall t_i \in [\mathrm{TD}_g + t_1, t_h] \cap \mathcal{T}_h. \tag{58}$$

These constraints must be modified for initial conditions as follows:

$$u_{gt} = u_g^0, \quad \forall g, t \in [t_1, \mathrm{TU}_g^R + \mathrm{TD}_g^R + t_1] \cap \mathcal{T}_h, \tag{59}$$

where we have the definitions

$$\mathrm{TU}_g^R := \max\left\{0, (\mathrm{TU}_g - \mathrm{TU}_g^0)u_g^0\right\}, \quad \forall g \in \mathcal{G},$$

$$\mathrm{TD}_g^R := \max\left\{0, (\mathrm{TD}_g - \mathrm{TD}_g^0)(1 - u_g^0)\right\}, \quad \forall g \in \mathcal{G},$$

$$u_g^0 := \begin{cases} 0 & \text{if } \mathrm{TD}_g^0 > 0 \\ & \forall g \in \mathcal{G}. \\ 1 & \text{if } \mathrm{TU}_g^0 > 0 \end{cases}$$

4.3.3.3 Ramping Limits

Ramping limits can be expressed straightforwardly as

$$p_{gt} - p_{g,t-1} \leq t_{\mathrm{int}} \mathrm{RU}_g, \quad \forall g, t \in \mathcal{T}_h \setminus \{t_1\}, \tag{60}$$

$$-p_{gt} + p_{g,t-1} \leq t_{\mathrm{int}} \mathrm{RD}_g, \quad \forall g, t \in \mathcal{T}_h \setminus \{t_1\}. \tag{61}$$

Initial ramping constraints are

$$\begin{cases} p_g^0 - t_{\mathrm{int}} \mathrm{RD}_g \leq p_{gt_1} \leq t_{\mathrm{int}} \mathrm{RU}_g + p_g^0 & \text{if } \mathrm{TU}_g > 0 \\ 0 \leq p_{gt_1} \leq \mathrm{SU}_g & \text{if } \mathrm{TD}_g > 0. \end{cases} \tag{62}$$

We need not include initial ramp up/down conditions here since they are covered by the generation limits presented next.

4.3.3.4 Generation Limits

Generation limit constraint formulations depend on the generator's minimum up time requirement. If $\mathrm{TU}_g = 1$,

$$p_{gt_i} \leq (\overline{P}_g - \underline{P}_g)u_{gt_i} - (\overline{P}_g - \mathrm{SU}_g)v_{gt} \quad \forall t, \tag{63}$$

$$p_{gt} \leq (\overline{P}_g - \underline{P}_g)u_{gt} - (\overline{P}_g - \mathrm{SD}_g)w_{g,t+1} \quad \forall t. \tag{64}$$

For generators with $\mathrm{TU}_g \geq 2$,

$$\begin{aligned} p_{gt} \leq (\overline{P}_g - \underline{P}_g)u_{gt} - (\overline{P}_g - \mathrm{SU}_g)v_{gt} \\ - (\overline{P}_g - \mathrm{SD}_g)w_{g,t+1} \quad \forall t. \end{aligned} \tag{65}$$

For generator g, if we have $u_{gt^+} = u_{gt^-} = 1$, then the bounds on the generation output p_{gt} are as follows:

$$p_{gt} \geq \max\{\underline{P}_g, -\mathrm{RD}_g t_{\mathrm{int}} + p_{gt_p}\}, \tag{66}$$

$$p_{gt} \leq \min\{\mathrm{RU}_g t_{\mathrm{int}} + p_{gt_p}, \overline{P}_g - 2\hat{r}^{\mathrm{OR}}_{gt} t_{\mathrm{int}}\}. \tag{67}$$

If $\overline{P}_g - 2\hat{r}^{\mathrm{OR}}_{gt} t_{\mathrm{int}}$ is less than the lower limit of p_{gt}, then we relax $\hat{r}^{\mathrm{OR}}_{gt}$ as follows:

$$\hat{r}^{\mathrm{OR}}_{gt} \leftarrow \overline{P}_g - 2\hat{r}^{\mathrm{OR}}_{gt} t_{\mathrm{int}}.$$

4.3.3.5 Reserve Constraints

We will consider it impossible to provide reserve during slow startup and shutdown processes. Thus we have the constraint

$$\begin{aligned} p_{gt_i} + r^{\mathrm{SR}}_{gt_i} + r^{\mathrm{NSR}}_{gt_i} + r^{\mathrm{OR}}_{gt_i} \leq \left(\overline{P}_g - \underline{P}_g\right) \\ \times \left(1 - \sum_{j=1}^{n_{\mathrm{SS}}-1} v_{gt_{i+j}} - \sum_{j=0}^{n_{\mathrm{SS}}-2} w_{gt_{i-j}}\right) \forall t_i \in \mathcal{T}_h. \end{aligned} \tag{68}$$

Further, we restrict the amount of each reserve product individually per their definitions as follows:

$$r^{\mathrm{SR}}_{gt_i} \leq \frac{\mathrm{RU}_g}{6} u_{gt_i}, \quad \forall t_i \in \mathcal{T}_h, \tag{69}$$

$$r^{\mathrm{NSR}}_{gt_i} \leq \frac{\mathrm{SU}_g}{6}(1 - u_{gt_i}) + \frac{\mathrm{RU}_g}{6} u_{gt_i} \quad \forall t_i \in \mathcal{T}_h, \tag{70}$$

$$r^{\mathrm{OR}}_{gt_i} \leq \frac{\mathrm{SU}_g}{2}(1 - u_{gt_i}) + \frac{\mathrm{RU}_g}{2} u_{gt_i} \quad \forall t_i \in \mathcal{T}_h. \tag{71}$$

4.3.3.6 Controllable Demand

Some loads have a controllable part. Let the demand from load k at time t be represented by

$$d_{kt_i} := d^{\mathrm{F}}_{kt_i} + d^{\mathrm{C}}_{kt_i}, \tag{72}$$

where $d^{\mathrm{F}}_{kt_i}$ is the fixed part of the load and $d^{\mathrm{C}}_{kt_i}$ is the controllable part of the load. The controllable part is bounded by

$$0 \leq d^{\mathrm{C}}_{kt_i} \leq d^{\max}_{k}. \tag{73}$$

For fixed loads, we have $d^{\mathrm{C}}_{kt_i} = 0$ for all t_i.

4.3.3.7 Renewable Generation

For renewable generation, we assume that they are fast-moving and thus do not have the ramping constraints of the thermal generators. Further, their model is very similar to that of the controllable demands. Renewable generation output at bus k is constrained by

$$0 \leq p^{\mathrm{R}}_{kt_i} \leq p^{\mathrm{Rmax}}_{kt_i}, \tag{74}$$

where $p^{\mathrm{Rmax}}_{kt_i}$ is the renewable energy capability at bus k at time t_i according to the forecast given at time t_0.

4.3.3.8 Grid Constraints

First, we formulate the injection constraint at each bus as

$$P^{\text{net}}_{k,t_i} := \sum_{g \in \mathcal{G}_k} \Big(\underline{P}_g \big(u_{gt_i} + \tilde{v}_{gt_i} + \tilde{w}_{gt_i} \big) + p_{gt_i} \Big) + p^{\text{R}}_{kt_i} - d_{kt_i}. \tag{75}$$

The transitional startup variables help represent the ramping up and down of generators and the corresponding power output (assumed to be linear) when starting up and shutting down.

From here, the network constraints can be included as (21) and (27) for the centralized RHUC or (29) – (30) and (31) for the decentralized RHUC.

4.3.4 Complete Problem

With these, a complete expression of the centralized RHUC problem at time t is

$$\begin{aligned} \min \quad & C(\mathbf{x}, t) \\ \text{s.t.} \quad & (68) \\ & (53) - (71), \, \forall g \in \mathcal{G} \\ & (73), \, \forall k \in \mathcal{L} \\ & (74), \, \forall i \in \mathcal{N} \\ & (21), \, \forall ij \in \mathcal{E} \\ & (27), \;\; (P^{\text{net}}_{i,t} \text{ defined as } (75)) \;, \, \forall i \in \mathcal{N} \end{aligned} \tag{76}$$

4.4 Economic Dispatch Formulation

4.4.1 Objective Function

The ED will fix all the on/off state variables, startup variables, and shutdown variables. Also, just one time step will be considered. Therefore the cost function of each generator will be the sum of the variable part of the cost (depending on the power set point) and the fixed cost for each generator. The total generator cost is represented

by

$$C_g(\mathbf{x}_g, t) = C_g^{\mathrm{Q}} p_{gt}^2 + C_g^{\mathrm{LV}} p_{gt} + C_{\mathrm{FV}}(g,t) u_{gt}. \tag{77}$$

The fixed cost for each generator at each time t is

$$C_{\mathrm{FV}}(g,t) = C_g^{\mathrm{NL}}. \tag{78}$$

Note that the startup and shutdown costs are not included in the ED. Those costs are accounted for and considered in the rolling UC.

The next part of the objective function is a penalty term for deviating from the schedule prescribed by the UC solution. Since the load is assumed to follow a linear path between the hourly forecast points, the generation will ideally also follow a linear path between the UC schedules prior to and following the ED dispatch time period t. Thus, the setpoint for the ED dispatch in time t for generator g will be

$$p_{g,t}^{\mathrm{SP}} = p_{g,t^-} + \frac{t - t^-}{t^+ - t^-}(p_{g,t^+} - p_{g,t^-}), \tag{79}$$

where t^- and t^+ are the times of the nearest two UC decisions in the past and future, respectively, and p_{g,t^-} and p_{g,t^+} are the corresponding power set points at those times as determined by the UC. If we have $|u_{g,t^-} - u_{g,t^+}| = 1$, we can approximate the startup process of the generator as a linear increase or decrease in power. We can model this by putting generator g into a "startup" mode for the ED interval $[t^-, t^+]$. The startup processes are defined below.

The penalty cost term can therefore be represented as

$$\omega(p_{gt} - p_{gt}^{\mathrm{SP}})^2, \tag{80}$$

where ω is a scalar cost parameter that is the same across all generators. The value of ω should be chosen as somewhat larger than the largest quadratic cost coefficient in the generator set.

We also have a cost for load curtailment. This term is simply represented as

$$\sum_{k \in \mathcal{L}} \omega_d (d_k^{\max} - d_{kt}^C). \tag{81}$$

The cost term ω_d should be larger than the costs of deploying reserve, ω_{SR} and ω_{NSR}, discussed below.

There is a cost for deploying the reserve that was scheduled for the current period. That is, ED30, which occurs at 20 minutes and solves for the operations in the 30-40 minute range, should respect the reserve that was scheduled by ED20. Just as ED30 respects the ramp constraints in the 20-30 minute range, it should include a penalty for using the reserve in that period. This cost term is denoted $C_{R^-}(t)$ and is constrained in the following way:

$$C_{R^-}(t) \geq 0 \tag{82}$$

$$C_{R^-}(t) \geq \omega_R \left(r_{t^-}^{\text{SR}} + r_{t^-}^{\text{NSR}} - \left(\overline{P} - p_{gt} \right) \right) \tag{83}$$

The final part of the objective function is a cost term for deviating from the scheduled ten-minute spinning and nonspinning reserve. This can be represented as

$$C_{\text{R}}(\mathbf{x}, t) = \omega_{\text{SR}} \left(\hat{r}_t^{\text{SR}} - \sigma_t^{\text{SR}} \right)^2 + \omega_{\text{NSR}} \left(\hat{r}_t^{\text{NSR}} - \sigma_t^{\text{NSR}} \right)^2. \tag{84}$$

where we have the auxiliary constraints

$$\sigma_t^q = \sum_{g \in \mathcal{G}} r_{gt}^q. \tag{85}$$

The values of ω_{SR} and ω_{NSR}, which may be equal, should be large enough to discourage reserve from deviating for economic reasons. The intention is to only deviate from scheduled reserve (i.e. deploy reserve) when other resources are unavailable.

The total objective function is then the following quadratic cost:

$$\begin{aligned} C(\mathbf{x}, t) = \sum_{g \in \mathcal{G}} \left[C_g(\mathbf{x}_g, t) + \omega (p_{gt} - p_{gt}^{\text{SP}})^2 \right] \\ + C_{R^-}(t) + C_{\text{R}}(\mathbf{x}, t) + \sum_{k \in \mathcal{L}} \omega_d (d_k^{\max} - d_{kt}^{\text{C}}). \end{aligned} \tag{86}$$

4.4.2 Constraints

For generator g, if we have $u_{gt^+} = u_{gt^-} = 1$, then the bounds on the generation output p_{gt} are as follows:

$$p_{gt} \geq \max\{\underline{P}_g, -\mathrm{RD}_g t_{\mathrm{int}} + p_{gt_p}\}, \tag{87}$$

$$p_{gt} \leq \min\{\mathrm{RU}_g t_{\mathrm{int}} + p_{gt_p}, \overline{P}_g - 2\hat{r}_{gt}^{\mathrm{OR}} t_{\mathrm{int}}\}. \tag{88}$$

Note that the term $\hat{r}_{gt}^{\mathrm{OR}}$ is multiplied by $2t_{\mathrm{int}}$ because the operating reserve is defined as the amount that can be provided in 30 minutes, whereas we are concerned with a window of length t_{int} hours.

If generator g has $u_{gt^-} = u_{gt^+} = 0$, then we limit the power output of the generators to their nonspinning reserve capability as follows:

$$\begin{cases} p_{gt} + r_{gt}^{\mathrm{NSR}} \leq \mathrm{SU}_g/6 & \text{if } g \in \mathcal{G}^{\mathrm{SS}} \\ p_{gt} + r_{gt}^{\mathrm{NSR}} \leq \mathrm{SU}_g/2 & \text{if } g \in \mathcal{G}^{\mathrm{FS}}. \end{cases} \tag{89}$$

4.4.2.1 Startup and Shutdown Processes

A startup process applies to generator g if $u_{gt^+} - u_{gt^-} = v_{gt} = 1$. A shutdown process applies if $u_{gt^-} - u_{gt^+} = w_{gt} = 1$. In either case, the model of the generator g in the ED problem is altered in the interval $[t^-, t^+]$ as described in the following sections. Recall that the startup and shutdown capabilities are given by the following:

$$\mathrm{SU}_g = \max\{\underline{P}_g, \mathrm{RU}_g\}, \tag{90}$$

$$\mathrm{SD}_g = \max\{\underline{P}_g, \mathrm{RD}_g\}. \tag{91}$$

We distinguish between "slow-start" units and "fast-start" units. To be considered fast-start, a generator must have both $T_g^{\mathrm{CS}} < 1$ and $\mathrm{TU}_g \leq 1$. Similarly, to be considered "fast-stop", a generator must have both $T_g^{\mathrm{CS}} < 1$ and $\mathrm{TD}_g \leq 1$.

For the period that the unit is in startup or shutdown mode, we modify the reserve constraints. Note that this will influence the reserve cost term in the objective

function, thereby providing incentive for other generators to increase reserve offers to compensate for units starting up.

4.4.2.2 Slow-Start Startup and Shutdown Process

The bounds on p_{gt} are changed to

$$\underline{P}_g t_{\text{int}} + p_{gt_p} \leq p_{gt} \leq \text{SU}_g t_{\text{int}} + p_{gt_p}. \tag{92}$$

This handles the generation limit constraint and ensures that the power output of the generator at time t^+ will lie somewhere in the range $[\underline{P}_g, \text{SU}_g]$. Additionally, we add the reserve constraint

$$p_{gt} + r_{gt}^{\text{SR}} + r_{gt}^{\text{NSR}} \leq \text{SU}_g t_{\text{int}} + p_{gt_p} \tag{93}$$

to limit the amount of reserve that can be offered in the startup process.

For shutdown, the bounds on p_{gt} are changed to

$$p_{gt_p} - \text{SD}_g t_{\text{int}} \leq p_{gt} \leq p_{gt_p} - \underline{P}_g t_{\text{int}} \tag{94}$$

with the constraint $p_{gt} \geq 0$ still in force. We add the additional constraint for reserve

$$p_{gt} + r_{gt}^{\text{SR}} + r_{gt}^{\text{NSR}} \leq p_{gt_p} - \underline{P}_g t_{\text{int}}. \tag{95}$$

4.4.2.3 Fast-Start Startup and Shutdown Processes

We assume that fast-start units can reach their startup capability in 20 minutes. The bounds on p_{gt} during the fast startup process are therefore changed to

$$\underline{P}_g t_{\text{int}} + p_{gt_p} \leq p_{gt} \leq 3 \cdot \text{SU}_g t_{\text{int}} + p_{gt_p} \tag{96}$$

with the reserve constraint

$$p_{gt} + r_{gt}^{\text{SR}} + r_{gt}^{\text{NSR}} \leq 3 \cdot \text{SU}_g t_{\text{int}} + p_{gt_p}. \tag{97}$$

Therefore we have constraints governed by the cases

$$\begin{cases} (96) \text{ - } (97) & \text{if } p_{gt_p} < \underline{P}_g \\ (87) \text{ - } (88) & \text{if } p_{gt_p} \geq \underline{P}_g. \end{cases} \tag{98}$$

Note that once the generation exceeds the economic minimum, the unit is no longer in startup mode, but is in normal operation.

Similarly, we assume fast-stop units can fully shut down from their shutdown capability in 20 minutes. The bounds on p_{gt} during the fast shutdown process are changed to

$$p_{gt_p} - 3 \cdot \mathrm{SD}_g t_{\mathrm{int}} \leq p_{gt} \leq p_{gt_p} - \underline{P}_g t_{\mathrm{int}} \tag{99}$$

with the reserve constraint

$$p_{gt} + r_{gt}^{\mathrm{SR}} + r_{gt}^{\mathrm{NSR}} \leq p_{gt_p} - \underline{P}_g t_{\mathrm{int}}. \tag{100}$$

Therefore we have the constraints

$$\begin{cases} (99) \text{ - } (100) & \text{if } p_{gt_p} > 0 \\ p_{gt} = 0 & \text{if } p_{gt_p} = 0 \end{cases} \tag{101}$$

with the constraint $p_{gt} \geq 0$ still enforced. As with the startup process, once the unit reaches 0 output, it is no longer in shutdown mode.

4.4.2.4 Reserve Constraints

The ED will be able to dispatch SR and NSR adjusting for what was committed by rolling UC. The bounds on the SR and NSR products are dependent on the state of the generator. If the generator is online, i.e. $u_{gt^-} = u_{gt^+} = 1$, the bounds are

$$0 \leq r_{gt}^{q} \leq \frac{\mathrm{RU}_g}{6}, \forall q \in \{\mathrm{SR}, \mathrm{NSR}\} \tag{102}$$

along with the constraint

$$p_{gt} + r_{gt}^{\mathrm{SR}} + r_{gt}^{\mathrm{NSR}} \leq \overline{P}_g. \tag{103}$$

If the generator g is in a startup process, the bounds are

$$\begin{cases} 0 \leq r_{gt}^{q} \leq \left(\mathrm{SU}_g - \underline{P}_g\right)/6, \, \forall q \in \{\mathrm{SR}, \, \mathrm{NSR}\} & \text{if } g \in \mathcal{G}^{\mathrm{SS}} \\ 0 \leq r_{gt}^{q} \leq \left(3 \cdot \mathrm{SU}_g - \underline{P}_g\right)/6, \, \forall q \in \{\mathrm{SR}, \, \mathrm{NSR}\} & \text{if } g \in \mathcal{G}^{\mathrm{FS}}. \end{cases} \tag{104}$$

Similarly, if g is in a shutdown process, the bounds are

$$\begin{cases} 0 \leq r_{gt}^{q} \leq \left(\mathrm{SD}_g - \underline{P}_g\right)/6, \, \forall q \in \{\mathrm{SR}, \, \mathrm{NSR}\} & \text{if } g \in \mathcal{G}^{\mathrm{SSh}} \\ 0 \leq r_{gt}^{q} \leq \left(3 \cdot \mathrm{SD}_g - \underline{P}_g\right)/6, \, \forall q \in \{\mathrm{SR}, \, \mathrm{NSR}\} & \text{if } g \in \mathcal{G}^{\mathrm{FSh}}. \end{cases} \tag{105}$$

If the generator is offline, i.e. $u_{gt^-} = u_{gt^+} = 0$, the bounds are

$$r_{gt}^{\mathrm{SR}} = 0 \tag{106}$$

$$\begin{cases} 0 \leq r_{gt}^{\mathrm{NSR}} \leq \mathrm{SU}_g/6 & \text{if } g \in \mathcal{G}^{\mathrm{SS}} \\ 0 \leq r_{gt}^{\mathrm{NSR}} \leq \mathrm{SU}_g/2 & \text{if } g \in \mathcal{G}^{\mathrm{FS}} \end{cases} \tag{107}$$

OR is not considered by ED since it requires 30 minutes to deploy.

4.4.2.5 Controllable Demand

Some loads have a controllable part. Let the demand from load k at time t be represented by

$$d_{kt} := d_{kt}^{\mathrm{F}} + d_{kt}^{\mathrm{C}}, \tag{108}$$

where d_{kt}^{F} is the fixed part of the load and d_{kt}^{C} is the controllable part of the load. The controllable part is bounded by

$$0 \leq d_{kt}^{\mathrm{C}} \leq d_{k}^{\max}. \tag{109}$$

For fixed loads, we have $d_{k}^{\max} = 0$.

4.4.2.6 Renewable Generation

For renewable generation, we assume that they are fast-moving and thus do not have the ramping constraints of the thermal generators. Further, their model is very

similar to that of the controllable demands. Renewable generation output at bus i is constrained by

$$0 \leq p_{i,t}^{\mathrm{R}} \leq p_{i,t}^{\mathrm{Rmax}}, \tag{110}$$

where $p_{i,t}^{\mathrm{Rmax}}$ is the forecast for renewable energy capability at bus i at time t.

4.4.2.7 Grid Constraints

First, we formulate the injection constraint at each bus as

$$P_{i,t}^{\mathrm{net}} := \sum_{g \in \mathcal{G}_i} p_{gt} + p_{i,t}^{\mathrm{R}} - d_{i,t}. \tag{111}$$

The SR and NSR are allowed to contribute to balancing the demand in the system, although the costs ω_{SR} and ω_{NSR} should be large in the objective function.

The injections are tied to the bus phase angles, after DC power flow approximation, by

$$P_{i,t}^{\mathrm{net}} = \sum_{j \in \delta_i \bigcup \{i\}} B_{ij}\theta_{j,t}. \tag{112}$$

We also require the line limit constraint

$$-\overline{F}_{ij} \leq B_{ij}(\theta_{i,t} - \theta_{j,t}) \leq \overline{F}_{ij} \tag{113}$$

for all lines ij.

4.4.3 Complete Problem

With these, a complete expression of the centralized ED problem at time t is

$$
\begin{aligned}
\min \quad & C(\mathbf{x}, t) \\
\text{s.t.} \quad & (49) \\
& (87)-(88), (102)-(103), \forall g \in \mathcal{G} \text{ s.t. } u_{gt^-} = u_{gt^+} = 1 \\
& (89), (106)-(107), \forall g \in \mathcal{G} \text{ s.t. } u_{gt^-} = u_{gt^+} = 0 \\
& (92)-(93), (104), \forall g \in \mathcal{G}^{\mathrm{SS}} \text{ s.t. } v_{gt} = 1 \\
& (94)-(95), (105), \forall g \in \mathcal{G}^{\mathrm{SSh}} \text{ s.t. } w_{gt} = 1 \\
& (98), (104), \forall g \in \mathcal{G}^{\mathrm{FS}} \text{ s.t. } v_{gt} = 1 \\
& (101), (105), \forall g \in \mathcal{G}^{\mathrm{FSh}} \text{ s.t. } w_{gt} = 1 \\
& (109), \forall k \in \mathcal{L} \\
& (110), (112), \forall i \in \mathcal{N} \\
& (21), \forall ij \in \mathcal{E}.
\end{aligned}
\tag{114}
$$

4.5 Test Cases

Results are shown for two test systems. First, results for a 6-bus system will be shown and discussed. The small size of the system allows results to be interpreted more easily than larger systems, where generation commitment schedules cannot be easily examined. All test cases with accompanying data are available at [18].

The 6-bus system topology was based on the Garver 6-bus system available in the MATPOWER distribution [83]. Generator and load data were modified, however. The generator at bus 0 is a renewable generator with associated forecasts and schedule of output. Conventional generators were located at buses 1 and 2, with the bus 1 generator having quadratic cost and the bus 2 generator having piecewise linear cost. Buses 3, 4, and 5 contained loads. The bus 5 load was curtailable up to 30% of its maximum demand. It was assumed that all load forecasts were perfect.

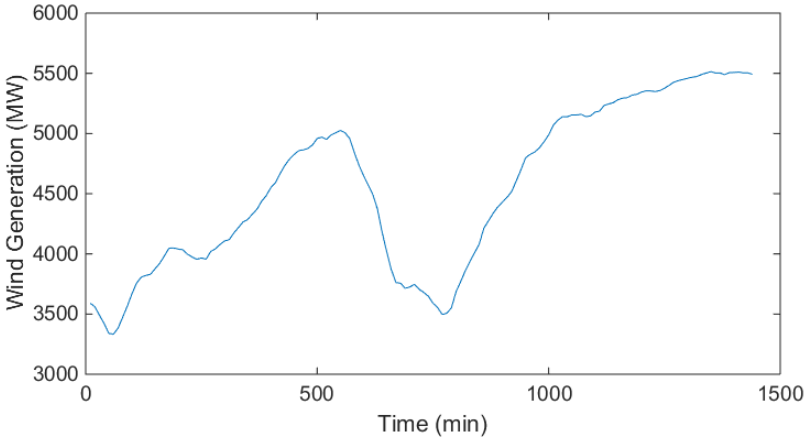

Figure 4.4: Total wind generation for 3,012-bus case.

The 3,012-bus case was based on the same case A_1 as examined in Chapter 3 but with two major modifications. First, of the 2,260 loads in the case, 253 of them were changed to flexible loads that were 50% curtailable, amounting to about 18.7% of the total load in each hour. This was done for the purpose of preserving feasibility of the case under various conditions and also to demonstrate the necessity to curtail load under some contingencies. Second, wind generators were added to 205 buses in the system. All wind generators were assumed to be fully curtailable.

The profiles of wind generation capability for each wind generator were estimated from wind data used in the Eastern Wind Integration and Transmission Study [23]. The data represented wind generator capability for every 10 minutes for one 24-hour period. The output was scaled to equal 20% of 3,012-bus case demand on average for the operating day. Figure 4.4 shows the wind generation throughout the operating day.

Two types of forecasts were generated for each wind case and site: a 24-hour forecast used in the 24-hour UC problem and 6-hour intra-day forecasts updated at each RHUC calculation interval. Forecasts were generated with error characteristics aligned with [8] and [23]. The 24-hour forecasts had a mean average error of 30%, while the 6-hour forecasts were near-perfect for the next hour but had a mean average

error of 15% for the remainder of the time horizon.

The 6-bus case models a scenario where, over 20 minutes, wind generation falls to 50% of the forecast capacity, recovering about an hour later. The 3,012-bus cases shown have two scenarios: one in which the forecast aggregated over the total system matches closely with output and one in which the wind generation falls to 67% of the forecast capability for about 40 minutes.

4.6 Experimental Results

We first show results for a 6-bus case to explain the meaning of the RHUC simulation outputs. We then show two 3,012-bus case scenarios to illustrate the implications of RHUC periodicity.

4.6.1 6-Bus Case

Figures 4.5 and 4.6 show the results of the rolling-horizon study on the 6-bus system with a wind outage using RHUC every 30 minutes and every 20 minutes, respectively. Solid blue lines represent total system generation as a result of the ED problem every 10 minutes while the dotted line segments represent total system generation as scheduled by each RHUC problem. In these two figures, each RHUC is annotated for clarification. Consider "UC_60" in Figure 4.5. This RHUC problem begins computation at 30 minutes with the information available at that time. UC_60 then solves for generation schedules at 70, 80, and 90 minutes, represented by the red dotted line. Output indicated by "UC_0" is generated here by the 24-hour UC. This and other dotted lines in Figure 4.5 are less visible since they lie directly on top of the blue generation output line. Overlap between RHUC and ED outputs occur when forecasts are perfect. The solid gray line represents curtailed load (no load was curtailed in either 6-bus case). Reserve usage (spinning and nonspinning reserve together) is shown in the dashed, black series.

The yellow dotted line represents the day-ahead forecast for wind generation while

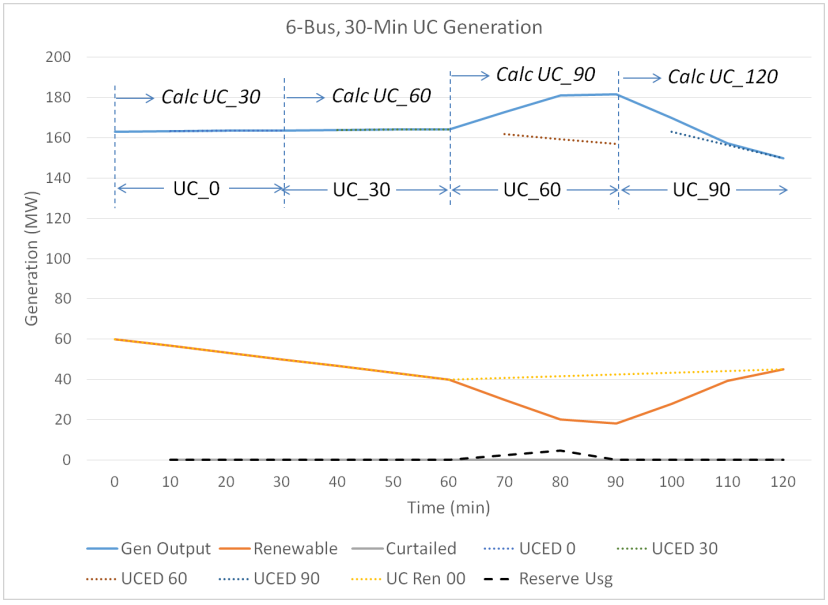

Figure 4.5: 6-bus system with moderate wind outage and 30-minute RHUC.

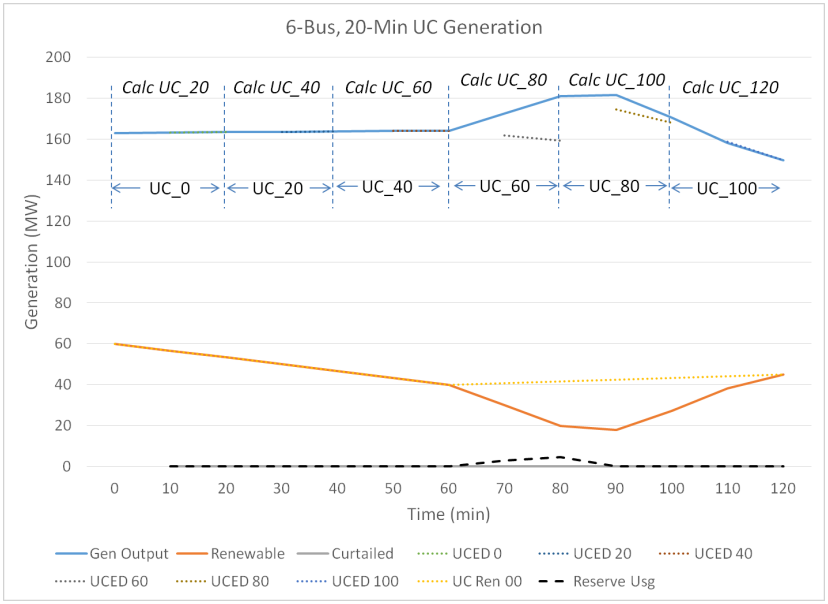

Figure 4.6: 6-bus system with moderate wind outage and 20-minute RHUC.

the orange solid line reflects the actual wind output from the ED solution. The updated wind forecasts for each RHUC are not shown. Deviation from forecast wind capability occurs from 70 minutes to 110 minutes with no advance warning. In other words, forecasts only reflect the information about wind deviation after it occurs. This makes the case effectively equivalent to an unplanned generator outage.

Forecasts for load and wind generation are perfect for the first 60 minutes of each 6-bus case, meaning RHUC generation schedules and wind output match perfectly with ED output series. However, after 60 minutes, the two scenarios deviate. Because the wind forecasts at 60 minutes remain aligned with the day-ahead wind forecast (wind does not drop out until 70 minutes), the 30-minute RHUC and the 20-minute RHUC output similar generation schedules for $t = 70$ min and $t = 80$ min. However, because in the 20-minute RHUC case another RHUC problem was solved at $t = 80$ min, generation output could be corrected in Figure 4.6 for $t = 90$ min and $t = 100$ min. This leads the ED solution to be closer to the RHUC solutions, which reduces the penalties incurred in the ED solution for deviating from the RHUC schedules. This is representative of the benefit of having the system follow more closely to an expected trajectory as in Figure 4.6 as opposed to Figure 4.5 where the system spends more time far away from the expected state.

In both the 20-minute and 30-minute RHUC cases, a small amount of reserve is deployed at $t = 70$ and $t = 80$ minutes. This lasts until generation can ramp up to balance demand at $t = 90$ minutes. The total amount of reserve available was about 14 MW, so the deployment was only a fraction of the capability.

We can also see this in Figure 4.7 where the generation costs for the 30-minute (solid red) and 20-minute (dashed yellow) RHUC scenarios overlap almost perfectly. Generation costs represent the sum of all terms represented by equation (77), while penalty costs represent the sum of scaled deviations from RHUC schedules as formulated in equation (80). As mentioned above, the difference between the two scenarios

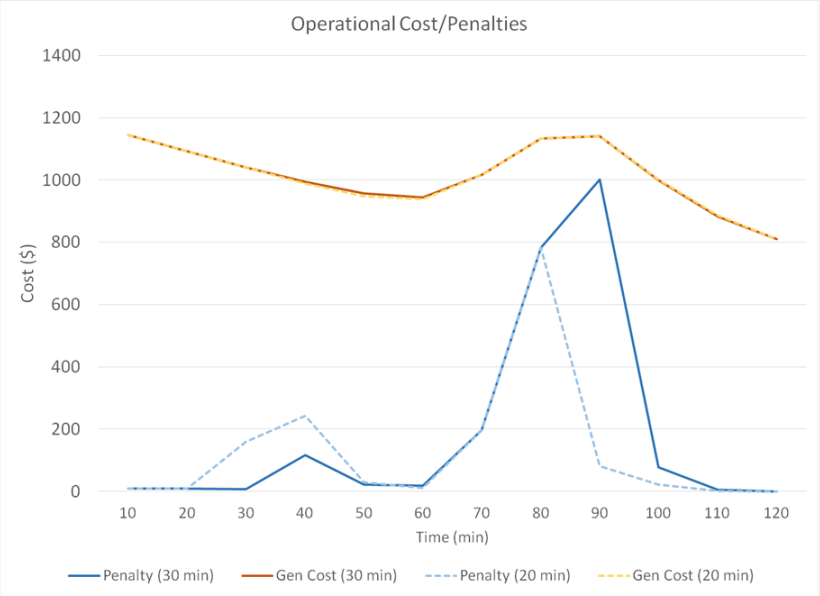

Figure 4.7: System cost & gen. penalties for 6-bus case with moderate outage.

is the deviation between the ED output and RHUC schedules, which is reflected in the larger penalties for the 30-minute RHUC scenario. Note that the penalties are nearly equal up to $t = 80$ min and that the 20-minute RHUC scenario yields much smaller generation penalties for the time frame [90, 100] minutes.

4.6.2 3,012-Bus Case

In Figures 4.8 and 4.9, we see the 3,012-bus case with expected wind output. Although each individual wind site had a forecast with mean average error of 30% in the day-ahead case, when aggregated among all the wind generators at the 205 "wind" buses, the day-ahead forecast was very close to the actual wind output values. The black, dashed line represents the scaled (down by a factor of 100) reserve penalty factor according to equation (84). The reserve usage for the 3,012-bus cases is scaled up by a factor of 10 to show more clearly its value on the graphs. No 3,012-bus case experienced load curtailment.

Results for these scenarios are largely as expected since the forecasts match the

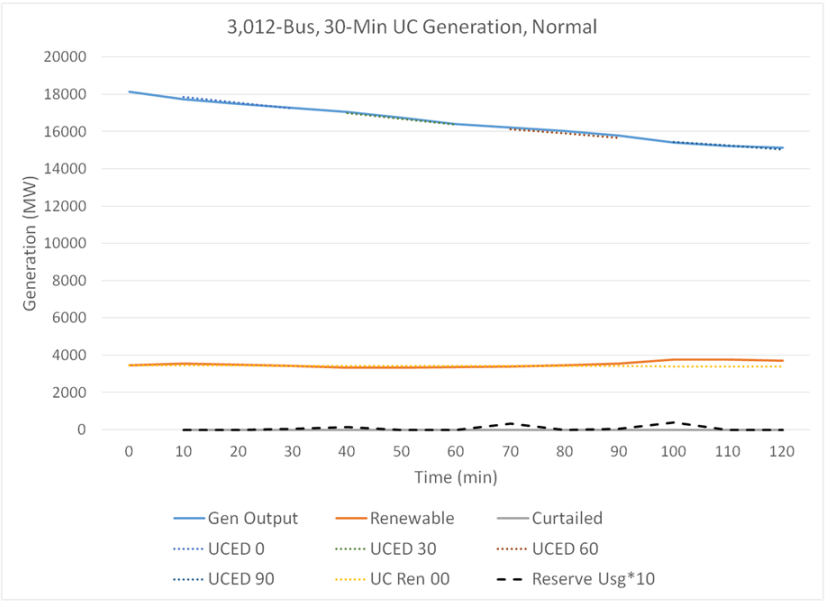

Figure 4.8: 3,012-Bus system in normal state with 30-minute RHUC.

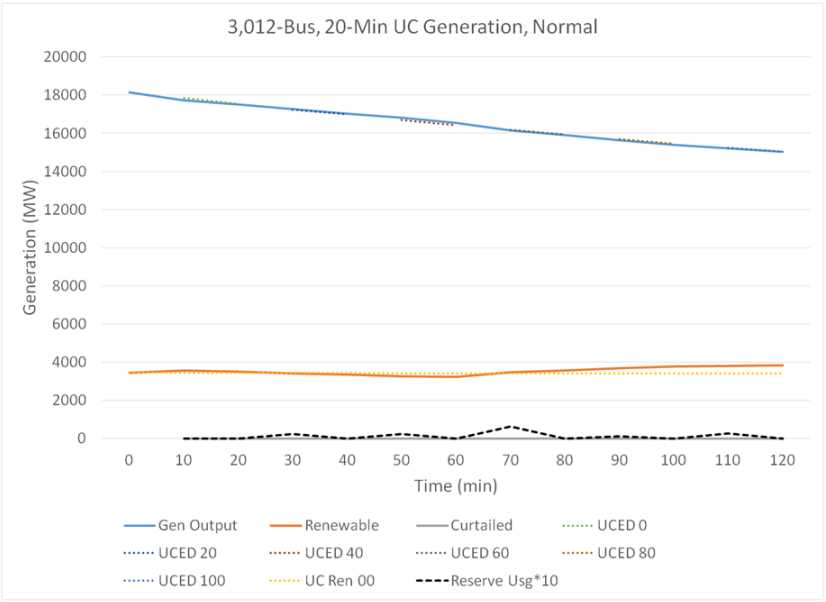

Figure 4.9: 3,012-Bus system in normal state with 20-minute RHUC.

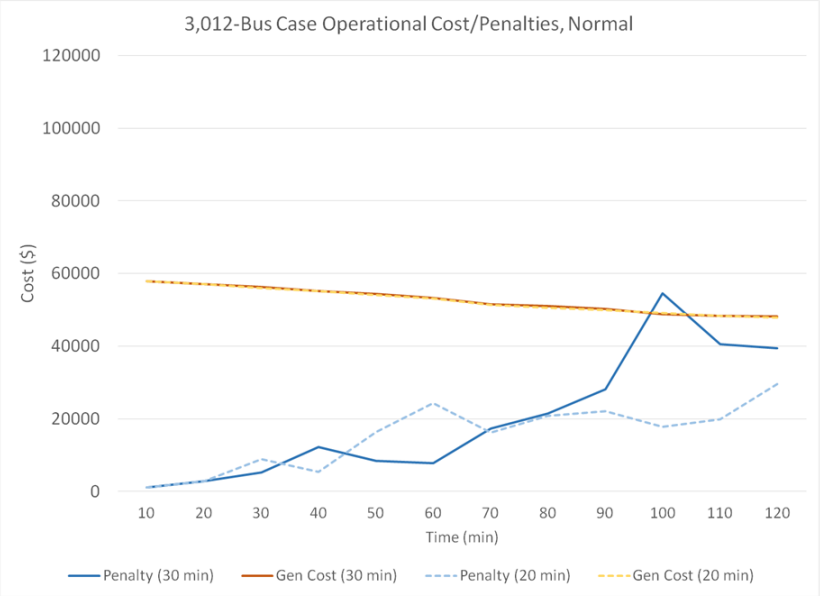

Figure 4.10: System cost & gen. penalties for 3,012-bus case in normal state.

observed conditions, meaning the RHUC and ED problems are not making many adjustments beyond the schedule created by the 24-hour UC problem. A small amount of reserve is deployed occasionally in both scenarios to manage the slightly fluctuating wind and smooth out some discontinuities in generator startups and shutdowns. These small reserve deployments coincide with the period of RHUC in each scenario because those very startups and shutdowns are scheduled according to the period of the RHUC.

Figure 4.10 shows that total generation costs for the 30- and 20-minute RHUC scenarios were nearly identical. Similar to Figure 4.7 for the 6-bus case, we notice that the 20-minute RHUC scenario results in smaller overall penalty terms than the 30-minute RHUC. Even for relatively small perturbations such as the wind output capability surplus observed from 90 to 120 minutes, conducting RHUC on a 20-minute period resulted in overall a better prediction of where the system would be headed than the 30-minute period.

The RHUC results for the 3,012-bus case with a 33% wind outage are shown in

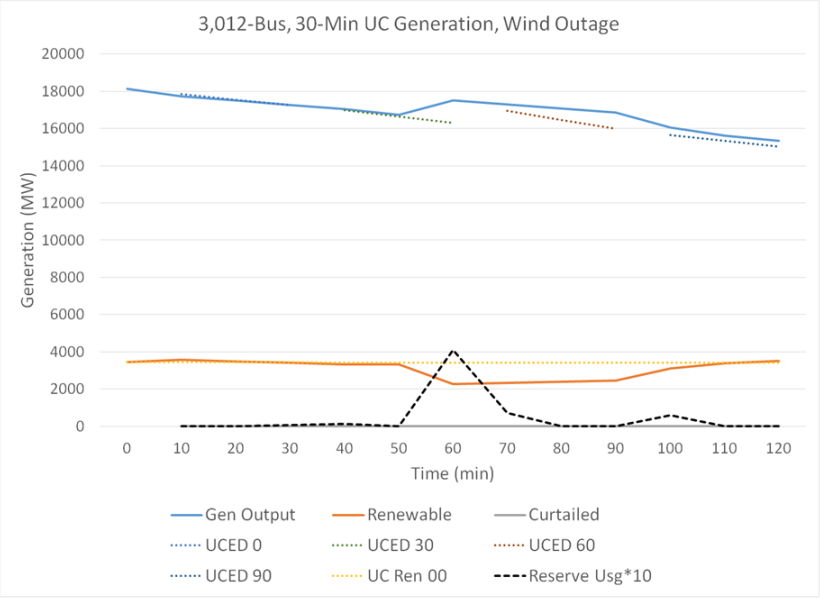

Figure 4.11: 3,012-bus system with wind outage and 30-minute RHUC.

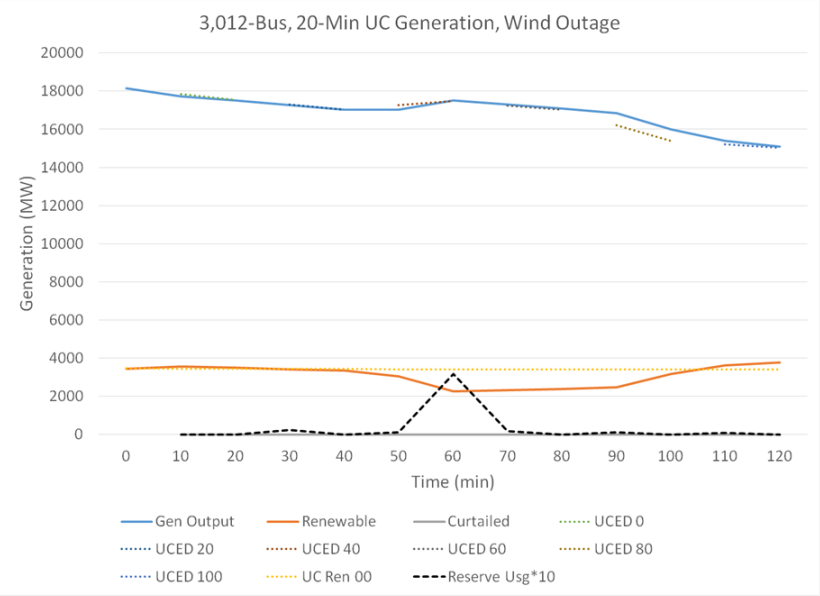

Figure 4.12: 3,012-bus system with wind outage and 20-minute RHUC.

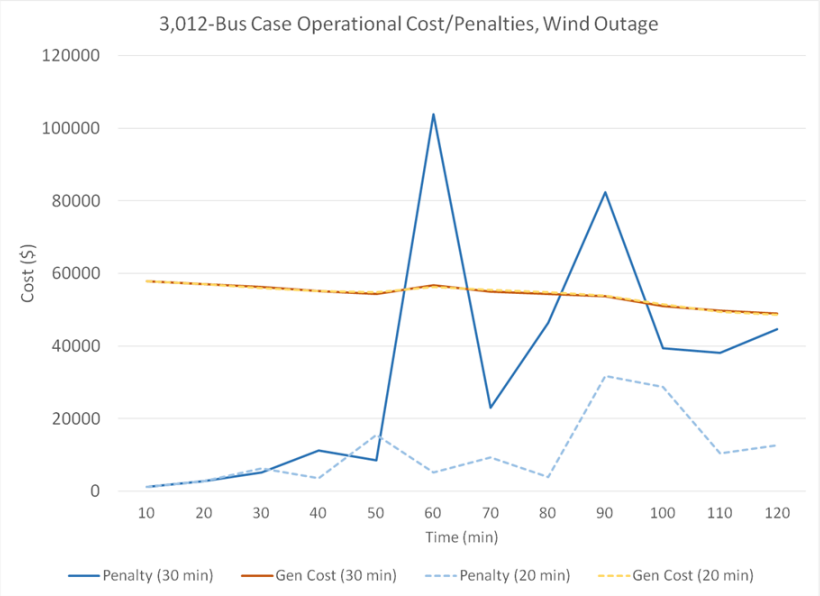

Figure 4.13: System cost & gen. penalties for 3,012-bus case with wind outage.

Figures 4.11 and 4.12. In this case, forecasting information for the outage approaching at 60 minutes becomes available at 40 minutes. The effect of this can be seen directly by comparing the two figures for the [50, 60] minute time frame. With 30-minute RHUC periodicity, the RHUC calculation for $t = 40$, 50, and 60 minutes must begin at $t = 30$ minutes, resulting in insufficient generation being scheduled and requiring about 100 MW more reserve to be deployed than with a 20-minute RHUC.

In contrast, the 20-minute RHUC periodicity allows the RHUC calculation to begin at 40 minutes when the new information is available, resulting in more committed generation and better-conditioned state for ramping up generators. The 20-minute RHUC yielded the full reserve requirement being scheduled at all times with the deployment of about 300 MW at the start of the wind outage event. Though it is not shown in Figure 4.11, with 30-minute RHUC, the full reserve requirement could not be scheduled at $t = 70$ minutes after the deployment of reserve at $t = 60$ minutes, so the system trajectory is even less secure in that scenario.

Figure 4.13 shows the generation cost and penalty terms for the 3,012-bus case

with wind outage. Again, cost of generation is very similar for the two cases even during the wind outage event from 60 to 90 minutes. Penalty terms were widely different for the two scenarios, though. In the 30-minute RHUC case, significantly more generation penalties were observed at $t = 60$ and $t = 90$ minutes than for the 20-minute RHUC case. This indicates that ED made major adjustments to generation outputs compared to the inherited RHUC schedule. While it is appropriate and necessary to make these adjustments, it can be dangerous to operate in a state far from that scheduled by a unit commitment operation. Decreased situational awareness, increased vulnerability to additional contingencies, and disrupted reserve allocations all make system operation very difficult under these circumstances. Therefore, the much lower penalties exhibited by the 20-minute RHUC scenario indicate a more desirable operational environment even though generation cost is similar to the 30-minute RHUC case. This benefit comes in addition to the lower deployment of reserve.

4.7 Conclusions

In this chapter, a framework for evaluation of rolling-horizon power grid operation was designed. Formulations were given for the RHUC problem with sub-hour time granularity and the ED problem inheriting from RHUC output. Results were demonstrated and discussed for a small, 6-bus system and a much larger 3,012-bus system under normal operation and a wind outage event scenario. RHUC periodicities of 20 and 30 minutes were compared. Although the generation costs were very similar in all cases between the 20- and 30-minute RHUC cases, the deviation of ED solutions from inherited RHUC schedules was noted to be superior for the 20-minute periodicity. Importantly, the 20-minute RHUC case overcame the wind outage event by deploying less of the reserve capacity, whereas the 30-minute RHUC case required more and was unable to subsequently recover the full reserve requirement to protect

against follow-on events. By developing and preparing rich datasets with historical forecasts or by using "back-casting" techniques with measured wind output data, many historical scenarios can be analyzed in view of this rolling-horizon framework to not only determine the benefits of faster RHUC periodicity, but also identify how close to optimal were the historical reserve deployment or load curtailment decisions.

CHAPTER V

CONTINGENCY CONSTRAINTS IN DUC

5.1 Introduction

5.1.1 Motivation

Transmission networks should be operated in a secure mode in which power is balanced with supply and line capacity constraints are not violated, as expressed in the prior formulations. However, the network should also be protected against collapse in the event of credible contingency scenarios. One type of contingency is the loss of a generator or, especially in the case of intermittent renewable generation, a drop in expected output. Minimum system reserve constraints are designed to mitigate the consequences of generator failures. If any generator in the system is suddenly taken offline, reserve units can ramp up within 10-30 minutes to make up the lost generation and restore system balance. This concern motivates the addition of reserve constraints of the types formulated in (17).

The other source of contingencies is loss of network elements. Network contingencies, which consist primarily of transmission line and transformer (i.e. branch) outages, change the topology of the system. Single-element contingencies, so-called $N-1$ contingencies, are the most common scenarios considered. The $N-1$ security criterion, in which the loss of a single element should not put the system in a compromised state, is the classical approach for guarding against network contingencies. Increasingly, operators are being required to consider $N-k$ contingencies. Although electric grids can rarely be fully $N-k$ compliant for $k > 2$, some credible contingencies for which $k = 2$ are often considered.

Network contingencies are more difficult to include in the UC problem than generator contingencies, which are generally considered to be satisfied through the reserve constraints. Therefore, network contingencies will be the focus of this chapter. Contingency constraints are an important part of secure generation scheduling. This chapter will propose two formulations that seek to include network contingencies in DUC modeling and solution.

5.1.2 Review of Unit Commitment Contingency Constraints

The set of contingencies explicitly considered in the optimization problem is much smaller than the entire list of credible contingencies. Only a subset of the credible contingency list is selected for actual modeling in the UC problem. Generally, by protecting against some severe contingencies, many other contingencies are also protected against. This can happen when the outage of a line has a symmetric effect on two other nearby lines.

Nonetheless, the process of contingency selection is important for obtaining a tractable model. Brute-force contingency analysis and selection requires the solution of many thousands of power flow scenarios and is very computationally burdensome for large systems. Therefore, many contingency selection methods focus on estimating the severity of contingencies to conduct deeper analysis on those most likely to threaten the system. Ejebe and Wolllenberg [22] provided an early, definitive performance index-based method for ranking and selecting contingencies. Updates to this ranking method, especially considering contingencies in which voltage limits are violated, were provided in [73]. For conducting the analysis of contingency scenarios, parallel computing approaches with dynamic load-balancing capability have been applied for both steady-state scenario analysis as in [41, 20] and transient analysis as in [46]. These selection techniques determine a list of contingencies that should be

passed to the subsequent operations stage. In this work, we do not focus on contingency selection or the process of contingency analysis. Instead, we will focus on how to include these constraints in DUC. Therefore, we will assume that a contingency list has already been established. Although [2] provided a way to further reduce this constraint set through identification of umbrella constraints, we will assume such a procedure has already been conducted and explicitly model all contingencies in the list.

The formulation of constraints related to line outages is given in Chapter 11 of [77] and will be reviewed below. In [56, 9, 74], which focus on the security-constrained optimal power flow (SCOPF) problem, contingency constraints are formulated in both pre-contingency and post-contingency action forms. The motivation behind post-contingency corrective actions is that small violations resulting from outages might be relieved by ramping up or down some generation or operating some other transmission device (e.g. tap changer or power electronic device) in a safe time frame. This approach works well for the SCOPF where a secure UC solution schedule has already protected against many of the most severe contingencies. However, UC generally considers only the most severe constraints and operates on a longer time horizon than SCOPF, so we will consider only preventive action in the precontingency mode. Recently, robust optimization theory has been applied to the network contingency-constrained UC problem in [58, 76]. However, applying these techniques in a decentralized setting is difficult because of the need to form dense Benders cuts in the subproblems and pass them back to the master problem. For this reason, we focus here on explicitly modeling each contingency in the selected scenario set.

Although Kargarian et. al. [45, 44] considered distributed UC problems, no explicit treatment of network contingency constraints was given. Because of the convexity of the SCOPF problem under the DC power flow approximation, distributed optimization methods have more often been applied to that problem than to the

UC. Most SCOPF literature seeking to apply distributed optimization or parallel computation focuses on decomposing the problem along each contingency scenario instead of regionally. Biskas and Bakirtzis [5] provided a decentralized formulation of the SCOPF problem, but only outages on or affecting tie lines of each subproblem were considered. More recently, [11] provided a regionally distributed formulation of SCOPF, applying ADMM and the proximal message passing algorithm to solve the problem. For network contingency constraints, a variable duplication technique was postulated similar to the one to be developed in Section 5.3.2, but performance of this strategy was not demonstrated since the only results presented were for a two-bus system. We contribute to the SCOPF literature by providing a new formulation for contingency constraints in Section 5.3.3 and comparing it to the variable duplication strategy for a 118-bus system and a 3,012-bus system. We contribute to the UC literature by including these constraints in distributed UC problems.

The remainder of the chapter proceeds as follows. In Section 5.2, we review the conventional formulation of contingency constraints for UC as well as the calculation of the line outage distribution factors (LODFs). Section 5.3 presents two alternative approaches for handling contingency constraints in the precontingency mode. The fully decentralized variable duplication strategy is discussed in Section 5.3.2. A less decentralized but still distributed (that is, computationally parallelizable) approximate approach is presented in Section 5.3.3. Results of simulations for varying levels of decomposition of a 118-bus system and a 3,012-bus system are presented in Section 5.4, and Section 5.5 concludes with a discussion of the two formulations, implications of contributions, and further work needed.

5.2 Conventional Contingency Constraint Formulation

Network constraints related to line contingencies can be formulated using linear sensitivities similar to (21). Instead of the sensitivity of a particular branch to the

injections at each bus, we consider the sensitivity of a branch to the flows on other branches. Let $\sigma_{ij,mn}$ denote the fraction of power from opened line mn that appears on line ij after the contingency. This is referred to as a line outage distribution factor (LODF). The conventional form of network contingency constraints based on such sensitivities is

$$-\overline{F}_{ij} \leq F_{ij,t} + \sum_{mn \in L^s} \sigma_{ij,mn} F_{mn,t} \leq \overline{F}_{ij}, \; \forall t, \; \forall s \tag{115}$$

where scenario s denotes the outage of line mn. Such constraints may be generalized to multiple-outage scenarios, but only single-outage scenarios will be considered here.

The LODFs for scenario s can be calculated according to [37] as

$$\mathbf{L}^s = \mathbf{\Gamma} \mathbf{C}^s \left(\mathbf{I} - \mathbf{\Gamma}^s \mathbf{C}^s \right)^{-1} \tag{116}$$

where $\mathbf{C}^s$ is the bus-to-branch incidence matrix for the outaged branches in scenario s and $\mathbf{\Gamma}^s$ is the GSF matrix for the same outaged branches. Note that for single or double line contingencies, the matrix inversions are trivial. The GSF matrix $\mathbf{\Gamma}^s$ can be calculated simply by

$$\mathbf{\Gamma}^s = \text{diag}(\mathbf{b}^s) \mathbf{C}^{sT} \mathbf{B}^{-1}, \tag{117}$$

where $\mathbf{b}^s$ is the vector of susceptances for the outaged branches.

The conventional UC formulation including contingency constraints then looks similar to (24) except with constraints (115) added:

$$\begin{aligned} \min \quad & \sum_{\nu \in \mathcal{P}} C_\nu(\mathbf{x}_\nu) \\ \text{s.t.} \quad & (2) - (16), \quad \forall g \in \mathcal{G}_i, \forall i \in \mathcal{N}_\nu, \forall \nu \in \mathcal{P} \\ & (17), (18) \\ & (20), \; (F_{ij,t} \text{ defined as } (21) \text{ and } (22)) \; , \forall ij \in \mathcal{E}' \\ & (115), \forall s \in \mathcal{S}, \forall ij \in \mathcal{E}' \end{aligned} \tag{118}$$

5.3 Contingency Constraints for DUC

5.3.1 Challenges

The LODFs must be calculated for each branch and for each contingency. Note that, because the inverse of the system admittance matrix $\mathbf{B}$ is required, direct calculation of the matrix of LODFs needs a centralized entity with global knowledge of the transmission network. Without using LODFs, the prosumers require some understanding of how the system state is altered under changes to the grid topology. This is challenging because the outage of a line has effects on branch flows that extend beyond the prosumers neighboring the one containing the outage.

In this section, two formulations of DUC are provided that protect against network contingencies. One retains the fully decentralized aspect of the approach in Chapter 3. The other assumes that a centralized calculation of LODFs is done before the optimization problem begins solving and from there remains a decentralized computation.

5.3.2 Formulation with Duplicated Variables

Note that the LODFs help provide an estimate of the approximate system state after a contingency given a precontingency state. In this formulation, we assume the LODFs are not available, so the postcontingency states must be determined in some other way. Here we will explicitly model each postcontingency state by adding a new set of voltage phase angle variables θ for each contingency scenario s. We call $\theta^s_{i,t}$ at time t the voltage phase angle at bus i following the event of contingency scenario s. Similarly, because the network topology has changed, we have B^s_{ij}, which is the susceptance of line ij following contingency scenario s. We will have $B^s_{ij} = B_{ij}$ if ij is not outaged in scenario s and $B^s_{ij} = 0$ if ij is outaged under s.

Postcontingency branch flows are defined as

$$F^s_{ij,t} := B^s_{ij}(\theta^s_{i,t} - \theta^s_{j,t}),\ \forall t,\ \forall s. \tag{119}$$

Branch capacity limits in the postcontingency states may be relaxed some above the regular limits. Usually, contingencies trigger remedial actions by system operators to resolve the overloaded branches quickly before the thermal loading results in other failures. We denote this constant F_{ELim} and gave it the value 1.15 for all simulations. The postcontingency line limits can then be represented as one of the following depending on whether the line's terminal buses are both internal (constraint (120)), one internal and one boundary (constraint (121)), or one boundary and one foreign (constraint (122)).

$$-F_{\mathrm{ELim}}\overline{F}_{ij} \le B^s_{ij}(\theta^s_{i,t} - \theta^s_{j,t}) \le F_{\mathrm{ELim}}\overline{F}_{ij},\ \forall t,\ \forall s, \tag{120}$$

$$-F_{\mathrm{ELim}}\overline{F}_{ij} \le B^s_{ij}(\theta^s_{i,t} - \tilde{\theta}^s_{\nu,j,t}) \le F_{\mathrm{ELim}}\overline{F}_{ij},\ \forall t,\ \forall s, \tag{121}$$

$$-F_{\mathrm{ELim}}\overline{F}_{ij} \le B^s_{ij}(\tilde{\theta}^s_{\nu,i,t} - \tilde{\theta}^s_{\nu,j,t}) \le F_{\mathrm{ELim}}\overline{F}_{ij},\ \forall t, \forall s. \tag{122}$$

We use the same set of generator power output and startup variables under all contingencies to enforce the need to take precontingency action to protect the system from violations under each scenario. Therefore we can express the additional bus balance constraints similarly to (31):

$$P^{\mathrm{net}}_{i,t} = \sum_{\substack{j \in \delta_i \bigcup \{i\} \\ \text{s.t. } j \in \mathcal{N}^{\mathrm{IB}}_\nu}} B^s_{ij}\theta^s_{j,t} + \sum_{\substack{j \in \delta_i \bigcup \{i\} \\ \text{s.t. } j \in \mathcal{N}^{\mathrm{BB}}_\nu \bigcup \mathcal{N}^{\mathrm{FB}}_\nu}} B^s_{ij}\tilde{\theta}^s_{\nu,j,t},\ \forall t,\ \forall s. \tag{123}$$

Finally, we also have new linking constraints

$$\begin{aligned} \tilde{\theta}^s_{\nu,i,t} = \tilde{\theta}^s_{\nu',i,t}, \quad & \forall \nu, \nu' \in \mathcal{P},\ \forall t \in \mathcal{T}, \forall s \in \mathcal{S} \\ \text{s.t.} \quad & i \in \mathcal{N}^{\mathrm{BB}}_\nu \cap \mathcal{N}^{\mathrm{FB}}_{\nu'} \text{ or } i \in \mathcal{N}^{\mathrm{FB}}_\nu \cap \mathcal{N}^{\mathrm{BB}}_{\nu'}, \end{aligned} \tag{124}$$

as well as average-of-perception terms (125) and dual updates (126):

$$\bar{\tilde{\theta}}^{s,m}_{\nu,i,t} := \frac{\sum_{\nu' \in \Delta_i} \tilde{\theta}^{s,m}_{\nu',i,t}}{|\Delta_i|} \quad \forall i \in \mathcal{N}^{\mathrm{BB}}_\nu \cup \mathcal{N}^{\mathrm{FB}}_\nu,\ \forall t, \forall s, \tag{125}$$

$$\lambda^{s,m}_\nu := \lambda^{s,m-1}_\nu + \rho\left(\tilde{\theta}^{s,m}_\nu - \bar{\tilde{\theta}}^{s,m}_\nu\right), \tag{126}$$

The primal residuals associated with $\theta^s_{\nu,i,t}$ are $(\tilde{\theta}^{s,m}_{\nu,i,t} - \bar{\bar{\theta}}^{s,m}_{\nu,i,t})$ in iteration m. The dual residuals are $\rho(\bar{\bar{\theta}}^{s,m}_{\nu,i,t} - \bar{\bar{\theta}}^{s,m-1}_{\nu,i,t})$. Let the vectors $\bar{\bar{\theta}}^m_\nu$ and λ^m_ν include the corresponding terms $\theta^{s,m}_{\nu,i,t}$ and $\lambda^{s,m}_\nu$ for ease of notation.

The augmented Lagrangian formulation including these constraints is then

$$\begin{aligned} \min \quad & \sum_{\nu \in \mathcal{P}} \mathcal{L}_{\rho,\nu} \left(\mathbf{x}_\nu, \bar{\bar{\theta}}^m_\nu, \bar{\bar{F}}^m_\nu, \bar{\bar{r}}^m_\nu, \lambda^m_\nu, \mu^m_\nu, \omega^m_\nu \right) \\ \text{s.t.} \quad & \mathbf{x}_\nu \in S'_\nu, \, (17), \, \forall \nu \in \mathcal{P} \end{aligned} \tag{127}$$

where we define S'_ν as

$$S'_\nu := \left\{ \mathbf{x}_\nu : \left\{ \begin{array}{l} (2)-(16), \quad \forall g \in \mathcal{G}_i, \forall i \in \mathcal{N}_\nu \\ (31), \; (P^{\text{net}}_{i,t} \text{ defined as } (22)) \; , \forall i \in \mathcal{N}_\nu \\ (26), \, (120), \, \forall ij \in \left\{ kl \in \mathcal{E} : k, l \in \mathcal{N}^{\text{IB}}_\nu \right\} \\ (29), \, (121), \, \forall ij \in \left\{ kl \in \mathcal{E} : k \in \mathcal{N}^{\text{IB}}_\nu, l \in \mathcal{N}^{\text{BB}}_\nu \right\} \\ (30), \, (122), \, \forall ij \in \left\{ kl \in \mathcal{E} : k \in \mathcal{N}^{\text{BB}}_\nu, l \in \mathcal{N}^{\text{FB}}_\nu \right\} \end{array} \right\} \right. \tag{128}$$

.

The number of constraints in (127) is increased over the number of constraints in (42) by $|\mathcal{S}|$ times the number of constraints in (25) – (27). Further, the number of variables in (127) is $|\mathcal{N}^{\text{IB}}| \times |\mathcal{N}^{\text{BB}}| \times |\mathcal{S}|$ more than (42) because of the added θ^s variables for each scenario s.

5.3.3 Formulation with Reduced Constraint Set

Because formulation (127) includes many additional constraints and variables for each subproblem, performance may suffer due to increased solution time for the subproblems driven by the greatly increased number of quadratic terms in the objective function. Therefore we seek an alternative approach for comparison.

For prosumer ν, we can use the LODFs to represent contingency constraints for each internal line resulting from an outage of an internal or boundary line of prosumer ν. Thus, we can express the constraints for lines i–j in ν due to outage of lines m–n

in ν as

$$-F_{\text{ELim}}\overline{F}_{ij} \leq F_{ij,t} + \sigma_{ij,mn}F_{mn,t} \leq F_{\text{ELim}}\overline{F}_{ij},\ \forall t. \tag{129}$$

The number of security constraints of this type is therefore the number of lines in prosumer ν times the number of outages considered in prosumer ν. Note that we may drop any constraints for which the following condition holds:

$$\left|\overline{F}_{mn}\sigma_{ij,mn}\right| \leq (F_{\text{ELim}} - 1)\overline{F}_{ij}. \tag{130}$$

Consider that when a line outage occurs, the flow previously observed on the outaged line will appear primarily on branches nearby. Therefore, if the system is operated in such a way that branches near a contingent line can absorb the expected flow in the event of an outage, then it is expected that the rest of the system will be protected as well in most cases. This observation opens the possibility of extending the previous structure of Chapter 3 for convergence on boundary line flows between neighbors to also include convergence on flows of contingent lines. That is, for an outage of line mn in prosumer ν, we propose to enforce the contingency constraint (115) on monitored lines within ν as well as all $\nu' \in \mathcal{P}_\nu$. If mn is a boundary line of prosumers ν and ν', then (115) will be enforced on all monitored lines within each $\nu'' \in \mathcal{P}_\nu \cup \mathcal{P}_{\nu'}$.

Assume that there exists a monitored outage of a line mn internal to prosumer ν', a neighbor of ν. For this outage, the line flow variable $F_{mn,t}$ will be available to prosumer ν' but not to prosumer ν. Therefore, prosumer ν must substitute a perception variable $\tilde{F}_{\nu,mn,t}$ for the constraint. Similarly, the variable $F_{mn,t}$, which is internal to prosumer ν', will be renamed to $\tilde{F}_{\nu',mn,t}$. Now, the constraint (129) for internal lines ij of ν not under outage monitoring can be expressed as

$$\begin{aligned} -F_{\text{ELim}}\overline{F}_{ij} \leq F_{ij,t} + \sigma_{ij,mn}\tilde{F}_{\nu,mn,t} \leq F_{\text{ELim}}\overline{F}_{ij}, \\ \forall ij \in \mathcal{E}^{\text{I}}_\nu \setminus \mathcal{E}',\ \forall mn \in \bigcup_{\nu' \in \mathcal{P}_\nu} \mathcal{E}_{\nu'} \cap \mathcal{E}',\ \forall t. \end{aligned} \tag{131}$$

If ij is a boundary line or under outage monitoring, we use (132) instead:

$$-F_{\mathrm{ELim}}\overline{F}_{ij} \leq \tilde{F}_{\nu,ij,t} + \sigma_{ij,mn}\tilde{F}_{\nu,mn,t} \leq F_{\mathrm{ELim}}\overline{F}_{ij}, \quad \forall ij \in \mathcal{E}^{\mathrm{B}}_{\nu} \cup \mathcal{E}', \forall mn \bigcup_{\nu' \in \mathcal{P}_{\nu}} \mathcal{E}_{\nu'} \cap \mathcal{E}', \forall t. \tag{132}$$

The coupling constraints, then, take the form of

$$\tilde{F}_{\nu,mn,t} = F_{mn,t}, \forall \nu \in \mathcal{P}_{\nu'}, \forall mn \in \mathcal{E}' \cap \mathcal{E}^{\mathrm{I}}, \forall t \tag{133}$$

for internal lines, where $\mathcal{E}'$ is the set of lines monitored for outage and ν' is the prosumer containing line mn. For boundary lines on outage monitoring, there are two containing prosumers ν' and ν'', so we have

$$\tilde{F}_{\nu,mn,t} = F_{mn,t}, \forall \nu \in \mathcal{P}_{\nu'} \cup \mathcal{P}_{\nu''}, \forall mn \in \mathcal{E}' \cap \mathcal{E}^{\mathrm{B}}, \forall t. \tag{134}$$

The average perception of $F_{mn,t}$ in iteration k is then

$$\bar{\tilde{F}}^{i}_{\nu,mn,t} := \frac{\sum_{\nu' \in \mathcal{P}_{\nu'}} \tilde{F}^{i}_{\nu',mn,t}}{|\mathcal{P}_{\nu'}|} \tag{135}$$

for internal lines mn and

$$\bar{\tilde{F}}^{i}_{\nu,mn,t} := \frac{\sum_{\nu' \in \mathcal{P}_{\nu'} \cup \mathcal{P}_{\nu''}} \tilde{F}^{i}_{\nu',mn,t}}{|\mathcal{P}_{\nu'}| + |\mathcal{P}_{\nu''}|} \tag{136}$$

for boundary lines. The objective function terms to add have the form

$$\mu^{i}_{\nu}(\tilde{F}_{\nu} - \bar{\tilde{F}}^{i}_{\nu}) + (\rho/2)\|\tilde{F}_{\nu} - \bar{\tilde{F}}^{i}_{\nu}\|^2,$$

similar to the line flow terms in (38) but with $\tilde{F}_{\nu}$ and $\bar{\tilde{F}}^{i}_{\nu}$ now including the outaged line terms. Also note that, if mn is a boundary line, we have

$$\tilde{F}_{\nu,mn,t} = \tilde{F}_{\nu',mn,t} = F_{mn,t},$$

which is already part of the coupling constraint set. Therefore we do not need to include new variables for outages on the boundary lines. Dual variables are updated the same as (40), and the associated primal and dual residuals are updated as before.

This formulation can now be expressed as

$$\begin{aligned} \min \quad & \sum_{\nu \in \mathcal{P}} \mathcal{L}_{\rho,\nu} \left(\mathbf{x}_\nu, \bar{\bar{\theta}}^m_\nu, \bar{\bar{F}}^m_\nu, \bar{\bar{r}}^m_\nu, \lambda^m_\nu, \mu^m_\nu, \omega^m_\nu \right) \\ \text{s.t.} \quad & \mathbf{x}_\nu \in S''_\nu, \ (17), \ \forall \nu \in \mathcal{P} \end{aligned} \tag{137}$$

where we define S'_ν as

$$S''_\nu := \left\{ \mathbf{x}_\nu : \left\{ \begin{array}{l} (2)-(16), \quad \forall g \in \mathcal{G}_i, \forall i \in \mathcal{N}_\nu \\ (31), \ (P^{\text{net}}_{i,t} \text{ defined as } (22)) \ , \forall i \in \mathcal{N}_\nu \\ (26), \forall ij \in \left\{ kl \in \mathcal{E} : k,l \in \mathcal{N}^{\text{IB}}_\nu \right\} \\ (29), \forall ij \in \left\{ kl \in \mathcal{E} : k \in \mathcal{N}^{\text{IB}}_\nu, l \in \mathcal{N}^{\text{BB}}_\nu \right\} \\ (30), \forall ij \in \left\{ kl \in \mathcal{E} : k \in \mathcal{N}^{\text{BB}}_\nu, l \in \mathcal{N}^{\text{FB}}_\nu \right\} \\ (131)-(132) \end{array} \right\} \right. \tag{138}$$

.

Note that, although formulation (127) can be solved with essentially the same algorithm as (42) (see Section 3.3.2), the algorithm to solve (137) needs to be slightly modified. In particular, we need an additional communication step to account for contingent boundary lines according to the average perception (136) for such lines. This is reflected in step 3a. Steps that can be performed simultaneously by prosumers are assigned the same number with letters "a", "b", and "c" to distinguish them.

Step 1. *Initialize primal variables* $\mathbf{x}^0_\nu$ *and dual variables* λ^0_ν *and* μ^0_ν *for each region* $\nu \in \mathcal{P}$*; set* $m = 0$*.*

Step 2a. *Each region* ν *sends* $\tilde{\theta}^m_{\nu,i,t}$ *and* $\tilde{\theta}^m_{\nu,j,t}$ *for all time* t *and for all* $i \in \mathcal{N}^{BB}_\nu \cap \mathcal{N}^{FB}_{\nu'}$ *and* $j \in \mathcal{N}^{FB}_\nu \cap \mathcal{N}^{BB}_{\nu'}$ *to its neighbor* ν'*.*

Step 2b. *Each region* ν *sends* $r^{q,m}_{\nu,t}$ *for all time* t *and all products* q *to a designated region* ν^**.*

Step 2c. *For each region* ν' *a neighbor of* ν*,* ν *sends* $\tilde{F}^m_{\nu,mn,t}$ *for all time* t *and for all* $mn \in \mathcal{E}' \cap \mathcal{E}_{\nu'}$ *to its neighbor* ν'*. From region* ν'*, region* ν *receives* $\tilde{F}^m_{\nu',pq,t}$ *for all time* t *and for all* $pq \in \mathcal{E}' \cap \mathcal{E}_\nu$ *from its neighbor* ν'*.*

Step 3a. *Each region ν, for each contingent boundary line $mn \in \mathcal{E}' \cap \mathcal{E}_\nu^B$ having foreign region ν' opposite the boundary, sends all the perceptions $\tilde{F}_{\nu'',mn,t}$ for each $\nu'' \in \mathcal{P}_\nu$ to neighbor region ν'.*

Step 3b. *Region ν^* calculates $\bar{\bar{r}}_{\nu,t}^{q,m}$, the new reserve targets for each region by* (36) *and sends the corresponding targets out to each region.*

Step 4a. *Each region ν computes $\bar{\bar{\theta}}_{\nu,i,t}^m$, for all $i \in \mathcal{N}_\nu^{BB}$ and time t using* (35) *and sends it back to all regions $\nu' \in \Delta_i \backslash \{\nu\}$.*

Step 4b. *For each boundary line ij such that $i \in \mathcal{N}_\nu^{BB}$ and $j \in \mathcal{N}_\nu^{FB}$, region ν updates $\bar{\bar{F}}_{\nu,ij,t}^m$ from* (37).

Step 4c. *For each contingent line $mn \in \mathcal{E}' \cap \mathcal{E}_\nu^I$, region ν calculates $\bar{\bar{F}}_{\nu,mn,t}$ using* (135). *For each $mn \in \mathcal{E}' \cap \mathcal{E}_\nu^B$, region ν calculates $\bar{\bar{F}}_{\nu,mn,t}$ using* (136).

Step 5. *Each region ν updates its primal and dual residuals. Dual variables are updated using* (39) – (41).

Step 6. *For given primal and dual tolerances $\epsilon^{Pri} > 0$ and $\epsilon^{Dual} > 0$, if $m > 0$, $\|\alpha^m\| \leq \epsilon^{Pri}$ and $\|\beta^m\| \leq \epsilon^{Dual}$* STOP *and output $\mathbf{x}_\nu^m$ as optimal decision for each region ν; otherwise go to Step 6.*

Step 7. *Each region ν updates $\mathbf{x}_\nu^{m+1}$ by solving a self-contained problem of the following form:*

$$\mathbf{x}_\nu^{m+1} = \underset{\mathbf{x}_\nu \in S_\nu''}{\operatorname{argmin}} \ \mathcal{L}_{\rho,\nu}\left(\mathbf{x}_\nu, \bar{\bar{\theta}}_\nu^m, \bar{\bar{F}}_\nu^m, \bar{\bar{r}}_\nu^m, \lambda_\nu^m, \mu_\nu^m, \omega_\nu^m\right) \tag{139}$$

Set $m \leftarrow m+1$ and go to Step 2.

5.4 Experimental Results

5.4.1 Test Cases

Formulations (127) and (137) were implemented in the same software framework as used for (42). Two test cases were used: the IEEE 118-bus case and the 3,012-bus case presented in Chapter 3. The 118-bus case was the same dataset as used in [80]. Two minor changes were made to the data:

- Initial power output for each generator was changed to the minimum power output for generators that were initially on.
- MVA capacity for line 86-87 was increased to 650 MW since generator 39 has a maximum power output of 650 MW and is connected to the rest of the system only by line 86-87.

For the 118-bus case, reserve requirements were set differently than the requirement mentioned in Section 3.2.3. Instead of 10-minute reserve equal to the size of the largest generator capacity, the 10-minute requirement was set to 7% of the total load with half of that total required to be spinning reserve. The 30-minute requirement was set to 3% of the load. This change was made because of feasibility issues in peak load periods. All lines in the system were monitored for capacity violations per (20) and (21).

The 3,012-bus case was modified as well. Updates to this case included:

- Of the 2,260 loads in the case, 253 of them were changed to flexible loads that were 50% curtailable, amounting to about 18.7% of the total load in each hour (the same modification was made in Chapter 4).
- Increased capacities of lines 456-508, 1055-1990, 1140-1353, and 1371-1621 to 185, 130, 130, and 130 MW, respectively, to accommodate the full range of power outputs from generators tied to the system through those lines.

A small subset of lines experiencing congestion was selected for monitoring for the 3,012-bus case. These lines were chosen by running several UC problems and choosing lines with binding capacity constraints and large dual variables associated with the congestion cost. Table 5.1 shows this list (note that all bus numbers are indexed to 0). Including a significantly greater number of monitored lines than this resulted violating in computational memory limits because of the large number of bus injection variables needed to formulate each constraint.

A contingency list was created for each case by solving the centralized problem with its respective monitored line list but without any contingency constraints. The resulting dispatch schedule was tested for each of the 24 hours for each potential outage to determine the presence and severity of overloads of monitored lines. Candidate lines to include for contingency monitoring were then included in subsequent UC runs. If the subsequent run was infeasible, some contingency monitored lines were removed. Once the UC problem became feasible, the last contingency list tested was used.

Tables 5.2 and 5.3 show the contingency lists for both test cases as well as the number of hours that violations occurred in each post-contingency case and the maximum overload for each overloaded line. Recall that the emergency line capacity limits for this study were set to 115% of the normal line capacity limit. Since all the contingent lines in each case result in maximum overloads of over 15%, all listed contingencies result in binding contingency constraints. Table 5.4 shows the size of the centralized problem including constraints for contingencies listed in Tables 5.2 and 5.3 and time required to find a solution within 1% optimal for each test case. Since the 118-bus case had 6 contingencies and 179 monitored lines, contingency constraints accounted for $6 \times 179 \times 24 = 25,776$ constraints in the centralized problem. For the 3,012-bus case, we had $9 \times 44 \times 24 = 9,504$ contingency constraints.

After solving the 118-bus and 3,012-bus test cases both with no constraints and with the contingency list, each was partitioned into regions and solved using R&F with the modifications based on the two formulation described above. For the 118-bus case, decompositions into 10,15, 20, 30, and 40 regions were used. For the 3,012-bus case, decompositions into 50, 100, 150, and 200 were used. The same partitions of the 3,012-bus case as used in Chapter 3 were also used here. In all instances, a maximum of 26 computational nodes were used including one node as a coordinator and messenger for stopping criteria. Therefore, the 118-bus case with 15 regions used

Table 5.1: 3,012-Bus Case Monitored Lines

Lines (Bus # - Bus #)			
5-346	119-1509	1168-1910	1697-2062
13-16	143-2063	1188-1501	1910-1915
15-58	162-2206	1214-1290	2018-2034
16-438	169-2319	1269-1293	2151-2680
18-20	177-2385	1371-1621	2190-2290
20-449	253-260	1511-1584	2190-2680
21-458	382-717	1511-1672	2215-2331
25-483	448-641	1511-1814	2411-2561
29-497	625-670	1511-2145	2472-2495
56-845	1055-1990	1537-1815	2738-2745
100-1330	1159-1901	1579-1937	
118-1510	1167-2068	1584-1671	

Table 5.2: 118-Bus Centralized Case Security Violations

Outage Line	Overload Line	# Overload Times	Max Overload (%)
4-7	7-29	19	53.47
24-26	22-31	12	17.96
64-67	29-37	1	15.50
	46-68	17	55.46
	48-68	17	51.89
67-80	68-76	20	45.79
79-80	68-76	20	45.79
88-91	81-82	2	15.05

16 nodes, but the 40-region case used 26 nodes with each of the subproblem nodes initially responsible for either one or two subproblems in each iteration.

5.4.2 Duplicate Variable Formulation Results

Table 5.5 shows relevant solution metadata for the 118-bus case. The duplicated variables formulation performed very poorly for the 3,012-bus case, with solutions only converging when residual tolerances were relaxed considerably. When using stopping criteria similar to that used for DUC experiments in Chapter 3, this formulation did not complete the ADMM-CR (QP relaxation) stage within the time limit of 7,200 s. Therefore, we do not show results for the 3,012-bus case under the duplicate variables

Table 5.3: 3012-Bus Centralized Case Security Violations

Outage Line	Overload Line	# Overload Times	Max Overload (%)
118-1510	119-1509	22	27.86
162-2206	2472-2495	6	26.87
382-717	5-346	4	19.06
625-670	610-670	21	19.84
1159-1901	110-1453	2	18.88
	1712-1731	1	16.82
	1731-1901	3	24.80
1168-1910	1910-1915	5	18.63
1188-1501	1579-1936	1	17.72
1579-1937	1579-1936	1	18.25

Table 5.4: Centralized, Contingency-Constrained Problem Data

Case	# Bin. Vars.	# Cont. Vars.	# Constr.	Upper Bnd. ($ 10^6)	Lower Bnd. ($ 10^6)	Gap %	Soln. Time (s)
118	1,296	14,666	76,399	0.815	0.810	0.62	4.14
3,012	3,600	42,002	93,079	12.852	12.759	0.73	2,014.08

Table 5.5: 118-Bus Case Duplicated Variable Form. Solution Data

# Region	# Iter	t_{best} (s)	# Cycle	Cost ($ 10^6)	Gap (%)
10	1,174	2,710	17	0.81995	0.010
15	1,192	1,610	21	0.81917	0.009
20	1,156	811	17	0.81777	0.007
30	1,744	930	38	0.82222	0.013
40	1,619	836	30	0.81748	0.007

formulation.

For the 118-bus case, time to find the best solution decreased rapidly with increasing levels of decomposition. However, the break-even point is reached quickly with the new formulation. Although solution time was halved between the 10- and 15-region decompositions and again between the 15- and 20-region decompositions, no computational benefit was observed beyond this partitioning.

This effect is likely due to three factors. First, the number of iterations conducted was larger for the 30- and 40-region decompositions. Second, the maximum number of computational nodes was set to 26, so upon decomposing into more than 25 regions, each iteration requires at best the time required to solve the two easiest subproblems in series. This is as opposed to the case where each region is assigned its own subproblem, where each iteration requires the time needed to solve the hardest subproblem. Third, the total communication required between computational nodes in the system rises with increasing numbers of regions, and since many more messages must be passed with the variable duplication strategy as opposed to the no-contingency case, the break-even point of communication time penalty versus subproblem solution time benefit from small subproblems should intuitively be shifted back from that observed in the results of Section 3.4.

Table 5.6 contains the post-contingency violation data for the decompositions of the 118-bus case solved using R&F with the formulation (127). Rows that are italicized represent post-contingency violations of lines that were overloaded in the base case. That is, since the base case already has line capacity violations, the same violations remain in post-contingency cases. Therefore, the contingency constraints may be working correctly to mitigate other overloads while these remain. In terms of mitigating the violations resulting from line outages, the 10-, 15-, and 20-region decompositions performed well, removing nearly all the violations resulting from outages included on the contingency list. The 10-region instance still suffered from violations

Table 5.6: 118-Bus Duplicated Variable Form. Contingency Violations

# Region	Outage Line	Overload Line	# Overload Times	Max Overload (%)
10	4-7	7-29	6	16.36
		29-37	*5*	*25.22*
	64-67	*29-37*	*15*	*46.55*
15	4-7	*0-1*	*3*	*16.23*
		0-2	*23*	*78.58*
		29-37	*24*	*125.56*
	24-26	*0-2*	*23*	*82.35*
		7-29	*24*	*75.85*
		29-37	*24*	*139.55*
	64-67	*0-2*	*23*	*82.60*
		7-29	*24*	*79.89*
		29-37	*24*	*147.28*
	67-80	*0-2*	*23*	*82.58*
		7-29	*24*	*79.52*
		29-37	*24*	*140.52*
	79-80	*0-2*	*23*	*82.58*
		7-29	*24*	*79.52*
		29-37	*24*	*140.54*
	88-91	*0-2*	*23*	*82.57*
		7-29	*24*	*79.37*
		29-37	*24*	*137.52*
20	4-7	*29-37*	*22*	*41.89*
	24-26	*29-37*	*21*	*40.17*
	64-67	*7-29*	*2*	*18.24*
		29-37	*24*	*52.81*
	67-80	*7-29*	*3*	*17.68*
		29-37	*24*	*38.48*
	79-80	*7-29*	*2*	*17.66*
		29-37	*18*	*38.01*
	88-91	*7-29*	*2*	*17.68*
		29-37	*18*	*38.51*
30	4-7	7-29	7	15.93
	24-26	22-31	2	16.98
	64-67	46-68	1	30.30
		48-68	1	27.03
	67-80	68-76	1	18.64
	79-80	68-76	1	18.64
40	4-7	7-29	9	16.11
	64-67	29-37	3	17.81
		46-68	1	15.27

in 6 time steps of no more than 16.36%, a relatively minor overage compared to the emergency overage limit of 15%. The 30- and 40-region decompositions exhibited more severe remaining violations in their solutions, although they did not include any base case line violations as the other decompositions did. In all decompositions, the approach of formulation (127) did mitigate the violations caused by the contingencies in Table 5.2, decreasing the average number of violated time steps on monitored lines for each contingency from 18.0 in the unsecured case to between 0 and 2.2 and the maximum violation from 55.46% to 30.30% (not considering base case violations).

The post-contingency violation data for the 3,012-bus cases solved with formulation (127) are not given here because all decompositions yielded solutions with hundreds of base case violations. The significant base case violations, which were prevalent in these solutions as well as on a smaller scale in the 15- and 20-region decompositions of the 118-bus case, were not seen in the solutions obtained in Chapter 3. Their appearance, therefore, should be attributed to the greatly expanded number of Lagrangian terms in the objective functions of formulation (127) over (42). Although penalty terms can be adjusted to give more weight to base case constraints than contingency constraints, tuning the parameters for each decomposition is difficult. The algorithm can tend toward providing a solution that satisfies all the explicitly included post-contingency constraints while ignoring base case constraints, such as the case of the 20-region decomposition, or toward a solution satisfying all base case constraints while leaving some post-contingency constraints to remain, such as the 40-region decomposition. Clearly, the latter of these options is preferable since contingencies are unlikely by definition whereas the base case represents the actual operational plan.

5.4.3 Reduced Constraint Set Formulation Results

Tables 5.7 and 5.8 show results regarding the best solution found from using R&F on the formulation (137). For the 118-bus case, solutions very close to optimal were

Table 5.7: 118-Bus Case Reduced Constraint Form. Solution Data

# Region	# Iter	t_{best} (s)	# Cycle	Cost ($ 10^6)	Gap (%)
10	1,899	251	7	0.81361	0.002
15	1,073	206	11	0.81407	0.003
20	1,213	198	7	0.81386	0.003
30	1,298	178	14	0.81043	-0.002
40	1,232	160	12	0.81744	0.007

Table 5.8: 3,012-Bus Case Reduced Constraint Form. Solution Data

# Region	# Iter	t_{best} (s)	# Cycle	Cost ($ 10^6)	Gap (%)
50	872	3,747	47	11.337	-0.035
100	1,347	1,969	65	11.310	-0.273
150	1,434	460	44	11.326	-0.132
200	1,352	1,630	84	11.329	-0.106

found. The times to find the best solution were roughly in the range of 3-4 minutes, but a solution within 1% of optimal was found in about 1 minute for the 40-region decomposition and 75 s for the 30-region decomposition. For the 3,012-bus case, many cycles between ADMM-BIN+ and ADMM-BIN– were conducted as compared with Table 3.7. The time to find the best solution was less than the centralized solution time of 2,014 s for the 100-, 150-, and 200-region decompositions.

The optimality gap in Table 5.8 is determined by comparing to the lower bound in the second row of Table 5.4 less the cost of curtailed load, which was $1,417,697 with a penalty cost of $1,000/MWh. The load costs in this case served the purpose of ensuring that all dispatchable generation was used before curtailing any load. Because the value of this penalty is arbitrarily chosen, we remove it when comparing the generator costs. The majority of load curtailment occurred in the first time period as generators shifted to avoid contingency violations.

Note that the optimality gap is negative for all scenarios of the 3,012-bus case and one scenario of the 118-bus case. This is because the best solution found by R&F has a base objective function value (that is, the objective function without Lagrangian or penalty terms) less than the lower bound in the contingency-constrained centralized

Table 5.9: 118-Bus Reduced Constraint Contingency Violations

# Region	Outage Line	Overload Line	# Overload Times	Max Overload (%)
10	64-67	29-37	1	15.70
		46-68	10	15.08
15	24-26	22-31	11	16.69
	64-67	46-68	13	16.67
20	24-26	22-31	1	15.46
	64-67	46-68	9	16.63
30	4-7	7-29	13	15.04
	64-67	46-68	5	17.21
40	64-67	46-68	9	18.02

Table 5.10: 3012-Bus Reduced Constraint Contingency Violations Summary

# Region	# Post-Con. Violations	# Viol. Lines w/o Base Viol.	Avg. Viol. Times	Max Overload w/o Base Viol. (%)
1	66	66	7.33	27.86
50	16	7	2.00	19.31
100	25	7	2.56	21.29
150	7	7	3.00	27.60
200	13	4	2.22	19.40

solution. Because the residuals of the relaxed constraints do not reach exactly zero, there is a chance for some relaxed constraints to remain violated in a solution even though the residual tolerances are low. The negative optimality gap in these cases indicate that some line capacities are violated by small amounts in the presented solution and some violations remain for contingency cases. This does not necessarily mean that no constraints are violated in the cases where the optimality gap is positive. It should also be noted that no load was curtailed in the decentralized case since the relaxation of line capacity and contingency constraints allowed for generation schedules with small violations that do not require curtailment.

Table 5.9 shows the contingency violations for the various decompositions of the 118-bus case. Compared to the post-contingency violations shown in Table 5.2, many fewer violations occurred in the decentralized cases after including contingency constraints. Not only were fewer lines exposed to post-contingency violations, but those

lines that were violated spent fewer time steps in violated states. The average number of cumulative time steps each line spent in a post-contingency violated state decreased by an average of 93.6 out of 108 total violated time steps in the unsecured, centralized case. Moreover, the maximum overloads occurring in the post-contingency states were significantly depressed. Whereas Table 5.2 shows 4 of the 6 contingencies produced violations of 45% or more, the maximum violation found in all of the decentralized cases is 18.02%.

For the 118-bus case, the most secure solution results from the smallest decomposition into 10 regions. The remaining violations do not stem from a single source. Examining the 10-region case results in Table 5.2, it was observed that an outage of line $64-67$ produced overloads slightly greater than 15% in lines $29-37$ and $46-68$. Lines $64-67$ and $46-68$ are internal to prosumer 7 while line $29-37$ is internal to prosumer 4. In the 10-region decomposition, prosumer 4 does not border prosumer 7. Since line $64-67$ is an internal line, prosumer 4 includes no constraints on or perception of the flow $F_{64-67,t}$ in line $64-67$. Therefore the overload may be a result of a binding constraint on $\tilde{F}_{7,64-67,t}$ being left out of the problem.

Line $46-68$ presents a different case since it is an internal line to prosumer 7. Because each region includes the constraints on its outages internally, the overload in line $46-68$ cannot be a result of a missing binding constraint. Instead, the overage at several time steps is due to the minor disagreement between the calculated average perception $\bar{\tilde{F}}_{64-67,t}$ and the actual value $F_{64-67,t}$. These disagreements result in some violations when the actual system state is calculated post-solution even though the constraints are satisfied by the subproblem-feasible solutions. In fact, because the post-contingency violations are reduced so significantly in Table 5.9 from those in Table 5.2, it is likely that all the violations in Table 5.9 are due to this remaining disagreement and not any missing binding constraints. Should it be the case that such binding constraints were missing, their shadow costs would be minuscule given

the size of the overages.

Because of the number of violations seen in the 3,012-bus decomposition solutions, a summary of the results is given in Table 5.10. As in Section 5.4.2, solutions for these decompositions tended to exhibit some line capacity overloads in the base cases. These pre-contingency overloads subsequently appeared as post-contingency violations when in fact they had already been present. For this reason, we present not only the total post-contingency violations but also the number of violations after removing the base case violations. Similarly summarized results for the centralized solution (1 region) are also given for comparison.

Like the 118-bus case, post-contingency violations are reduced significantly from the unsecured, centralized case. The average number of time steps with violations (calculated by summing the number of violated time steps over the violated lines for each of the 9 outages) decreased from 7.33 in the centralized case to between 2 and 3. Many of the post-contingency violations (not resulting from base-case violations) in the decentralized scenarios were modest overages of 15-16% likely resulting from disagreement between $F_{ij,t}$ and $\bar{\bar{F}}_{ij,t}$ as discussed above. However, because of the choice of partitions and subsequent elimination of some constraints, several post-contingency violations were not removed by the neighbor-focused formulation of contingency constraints.

For example, in the 150-region decomposition, the overage of line $1579-1936$ due to outage $1188-1501$ retained a single-time-step overload of 18.94%, indicating that the corresponding violation indicated by the second-from-bottom row in Table 5.3 was not relieved. Nonetheless, many post-contingency violations that were not totally eliminated, such as overloads of line $610-670$ from outage $625-670$, which appeared in all four decompositions, were ameliorated. In that example, the number of overload times decreased from 21 in the unsecured centralized case to an average of 6.75 across the four decompositions.

5.5 Discussion and Conclusions

5.5.1 Discussion of Results

Comparing Table 5.5 to Table 5.7, we see that the reduced-constraints formulation (137) based on LODFs outperformed the duplicated-variables formulation (127) by providing solutions of similar objective function value in roughly an order of magnitude less solution time for a similar number of iterations. The objective function value was \$817,441 on average for the reduced-constraints formulation compared to \$819,318 for the duplicated-variables formulation. Further, comparing Table 5.6 to Table 5.9, we see that the reduced-constraints formulation results in no base-case violations and less severe post-contingency violations (a maximum violation of 18.02% as compared with 30.30%).

For the 3,012-bus case, we find a similar comparison. Although Table 5.10 reports some base-case violations, they are far fewer and less severe than the results of the duplicated-variables formulation. The reduced-constraints formulation had difficulty enforcing the last 10% of contingency violations and line capacity constraints even when residual tolerances were tightened, though the majority of violations were eliminated. By contrast, the duplicated-variables formulation was unable to arrive at a solution to the relaxed QP problem in the required time frame. The reduced-constraints formulation can therefore be said to scale much better with system size than the duplicated-variables formulation.

Neither of the formulations were able to completely remove all post-contingency violations and base-case violations simultaneously. Additional testing with more iterations in the ADMM-CR stage, more total iterations, and decreasing the total primal residual tolerances resulted in much the same solution set as attained previously (albeit with fewer total solutions in the case of decreasing residual tolerances). To improve security and feasibility of solutions from the reduced-constraints formulation,

more refined stopping criteria based on maximum outage line residuals should be investigated. To support this, independent penalty terms for each outage line may be necessary to module commitment decisions to enforce severe contingency constraints while lessening the impact of other contingencies in the objective function.

5.5.2 Links to Bounding Theory

Note that this approach is compatible with the "bounding" theory developed by Brandwajn [7]. According to this theory, we can confirm that the entire system (call this region N2) is secure to an outage of line mn if it is secure for a subregion N1 of N2 surrounding line mn and a specific criterion holds on the boundary region N3 surrounding N1. This criterion, which is elaborated in [77], can be stated as

$$\underset{\max}{\Delta} F_{pq,t} < b_{pq} \left| \underset{\max}{\Delta} \theta_{i,t} - \underset{\max}{\Delta} \theta_{j,t} \right|, \forall t, \tag{140}$$

where pq is a line in N2 and ij is a boundary line in N3. The term $|\Delta_{\max}\theta_{i,t} - \Delta_{\max}\theta_{j,t}|$ is the maximum change in voltage phase angle spread across the boundary region N3 as a result of the outage of line mn in N1. The term $\Delta_{\max}F_{pq,t}$ is equal to $|\overline{F}_{pq} - |F_{pq,t}||$, which represents the largest change in flow that can be allowed on pq post-contingency without causing a violation. The theorem, proved in [7], demonstrates that

$$\left|\Delta\theta_{p,t} - \Delta\theta_{q,t}\right| < \left| \underset{max}{\Delta} \theta_{i,t} - \underset{max}{\Delta} \theta_{j,t} \right|, \forall t,$$

for all pq outside of the boundary N3. We can use criterion (140) in a decentralized way following the DUC solution to determine whether a contingent line is still causing a violation somewhere in the system (and if so, by how much to modulate the flow on the contingent line to relieve the violation) using the following algorithm. For each prosumer ν:

Step 1. *Calculate* $|\Delta_{max}\theta_{i,t} - \Delta_{max}\theta_{j,t}|$ *on the boundary region and broadcast to all prosumers.*

Step 2. *Receive* $|\Delta_{max}\theta_{m,t} - \Delta_{max}\theta_{n,t}|$ *from each other prosumer (steps 1 and 2 amount to a "gather-all" operation).*

Step 3. *Determine whether, for any line pq in ν and any received value from step 2, criterion* (140) *is violated.*

Step 4. *If there is exists any violation in step 3, send*

$$\underset{max}{\Delta} F_{pq,t} - b_{pq} \left| \underset{max}{\Delta} \theta_{m,t} - \underset{max}{\Delta} \theta_{n,t} \right|$$

for all violated t to the relevant prosumers $\mathcal{P}_{viol}$ to notify of violation.

Step 5. *Each prosumer in $\mathcal{P}_{viol}$ checks for ability to redispatch to resolve violation in pq for times t while keeping all boundary branch flows within a tolerance ϵ. If possible, then STOP. If not possible, then another DUC is necessary.*

As mentioned by Galiana in [31], the requirement that a system be redispatched to resolve a contingency having violations far away from the outage (i.e. running a new DUC after step 5) is very infrequent.

5.5.3 Conclusions

Two formulations were presented and compared for introducing $N-1$ security-type constraints to the DUC problem. Both formulations extend work previously done in distributed OPF work and are compatible with existing methods of conducting contingency selection and analysis. However, the reduced-constraints formulation was found to be more effective overall than the duplicated-variables formulation because of the order-of-magnitude decrease in solution time and more reliable solution quality. Although neither method fully removed all post-contingency constraints and base-case violations for the test cases studied, they greatly ameliorated post-contingency constraints and reduced violations to within a few percent of the tolerable emergency limit. With some extensive tuning of stopping criteria based on different metrics than currently used and independent penalty parameters for contingencies, the outlook is good that the reduced-constraint formulation can be refined for use with very large-scale cases.

CHAPTER VI

CONCLUSIONS

6.1 Summary of Contributions

In this dissertation, decentralized unit commitment was motivated as a key power systems problem to solve in future grids that may be operated under a decentralized architecture. Appropriate formulations were derived to solve this problem and its extensions, experimental results demonstrated the performance of DUC, and its value was supported through a realistic operations framework. The contributions of this dissertation are as follows:

1. Formulation of system reserve constraints in DUC.

2. Experimental results demonstrating DUC solution time and solution optimality gaps within reasonable tolerances for large-scale power systems.

3. Formulations for rolling-horizon unit commitment problems as part of an interlinked operational framework conducted by power system operators.

4. Formulations for economic dispatch problem including penalties for schedule deviations and deployment of reserve.

5. Experimental results of rolling-horizon simulation platform to demonstrate potential value of conducting unit commitment and economic dispatch problems on various periodicities.

6. Extension of R&F to handle $N-1$ security constraints in fully decentralized way using a variable duplication approach.

7. A superior extension of R&F to handle network security constraints in a computationally distributed way using a reduced, linear sensitivity-based constraint set and regional bounding.

8. Motivation for decentralized power system operations in general and a framework for decentralized electric grid architectures.

A listing of the past and planned publications associated with this work is as follows:

- "Prosumer-based Control Architecture for the Future Electricity Grid," presented at the September 2011 IEEE Conference on Control Applications.

- "Large-Scale Decentralized Unit Commitment," under revision with the *International Journal of Electric Power and Energy Systems*.

- "Decentralized Security-Constrained Unit Commitment," planned based on Chapter 5 to submit to *IEEE Transactions on Power Systems* in Summer 2015.

- "Improved Rolling-Horizon Unit Commitment through Faster Periodicity," planned based on Chapter 4 to submit to *IEEE Transactions on Power Systems* in Summer 2015.

We now briefly summarize the contributions made from each of the elements.

6.2 Prosumer-Based Architecture

Many factors are contributing to increased stress in power system infrastructure. Centralized control architectures are now presenting limitations in several different layers of the grid, from device control to electricity markets. The new capabilities promised by recent developments in decentralized control, distributed computing, power electronics, and optimization theory require that the architecture of the grid be revisited.

To understand how to shape the grid in this new direction, a new framework is needed. The prosumer concept elaborated in Chapter 2 seeks to describe some of the most important elements of the future grid. Grid operations become prosumer interactions, from which a service-oriented structure emerges to support trade of energy products between prosumers. Communication systems therefore must become time-aware to give system agents real-time guarantees on when services will be delivered and to notify them promptly when conditions change or communication is lost. Although this structure places design constraints on communication systems that may be different from traditional designs, it does not necessarily require them to be far higher bandwidth or lower latency than they are today. In fact, by containing system complexity within prosumers and exchanging only necessary service information, it is hoped that communication systems can be simplified beyond what has been proposed for many microgrid systems.

When complexity is encapsulated in prosumers, though, different operations procedures are required. Central operators do not necessarily have global visibility of the system. Further, consumer-side generation means that previously unidirectional flows change direction to satisfy the service exchanges between prosumers. Therefore, decentralized and distributed system control is needed for several applications. This, along with the potential for computational speed gains, is the motivation for decentralized unit commitment.

6.3 Decentralized Unit Commitment with Contingency Constraints

To solve the UC problem in a decentralized way, a voltage-based, as opposed to the traditional injection-based, formulation was presented. An augmented Lagrangian method known as ADMM was applied to the problem because of its ability to exploit the separability of the regional problems and to overcome the need for strict convexity of the objective function. However, ADMM must be modified in the environment

of the UC problem because of the binary variables indicating unit on/off status. Therefore, the R&F method was developed to mitigate the oscillations in binary variables that occur in subsequent iterations of the subproblems. Experimental results showed this method to be successful in solving large-scale problems up to 3,012 buses much more quickly than a centralized solver could. This form of DUC can provide a way to clear electricity markets or to leverage parallel computing resources to solve UC more quickly than otherwise possible, a very useful proposition in planning studies wherein many UC problems must be solved representing potential future scenarios.

However, for reliability-focused UC problems, such as ISOs conduct following the day-ahead market-clearing UC, it is necessary to be able to include $N-1$ network security constraints. Two potential formulations for this objective were investigated. The first retains the fully decentralized structure of the original DUC but showed very poor performance, often resulting in large infeasibilities in the pre-contingency power system state even when contingency constraints were satisfied. The second compromised an aspect of decentralization by allowing an initial calculation of global linearized sensitivity factors. This enabled each prosumer to formulate contingency constraints using far fewer variables than the first formulation. Moreover, the constraints associated with each contingent element could be concentrated in the area of the associated prosumer, actually making the second formulation more scalable than the first. This second formulation provided much better performance and largely satisfied contingency constraints for the large, 3,012-bus case, although some violations did remain.

This work contributes a new formulation, a new solution method, and experimental results on large systems in the emerging research area of decentralized unit commitment. In the short term, DUC can be used by system operators to provide good, feasible solutions to difficult UC problems quickly in parallel with centralized solvers. Some UC problems are more difficult for centralized solvers to complete than

others, and as demonstrated in Section 3.4, DUC may find solutions much faster. To completely replace today's commercial UC solvers, some future work is necessary in addition to general commercial refinements. First, stopping criteria and the evaluation of feasibility versus each contingency constraint should be revisited. Fine-tuning maximum tolerable residuals can be done case-by-case, but more general rules are needed as the system changes topology regularly. A decomposition approach where contingency constraints are only added if violated may also help with scalability. Second, a method for improving the lower bound estimate is necessary to improve the optimality gap in a decentralized way. Empirically, the optimality gaps found through DUC are near industry tolerances (roughly 2% or less), but the best gap that can reported by DUC alone are much larger since the only lower bound available is from the solution to ADMM-CR.

Nonetheless, the R&F method provides a way to adapt ADMM to general MIP problems with a separable structure. This allows for an alternative approach to conventional branch-and-cut strategies which may not perform well for some problems. Part of the usefulness of R&F is that, since the subproblems are MIQP, it still allows for the modeling flexibility afforded by MIP structures barring the coupling constraints.

6.4 Rolling Horizon Unit Commitment

The computational speed benefit observed in the results of Section 3.4 provides one of the most intriguing short-term benefits for system operators. The ability to conduct UC more quickly allows it to be performed with greater frequency. UC is performed repeatedly throughout the operating day in order to alter commitment schedules in accordance with changes in system load, outages both planned and unplanned, and intermittent renewable generation.

Much contemporary literature focuses on using stochastic or robust optimization

approaches to solve ahead-of-time a UC problem that includes the elements of uncertainty in it and positions the system to protect against as many of them as possible or the worst-case of them. This is a good strategy, but it does not capture the system operator's ability to adjust to outages that were not predicted or modeled prior to the beginning of the UC calculation. Therefore, agility to respond to system events or new information about future events, which may only be available very soon before those events (such as can be the case with wind troughs), is also very important and an area of improvement.

Further, the ability of a system operator to procure reserve capacity and deploy it is a crucial part of their role. Stochastic or robust optimization approaches often diminish or eliminate the need for reserve because they model scenarios that the system should withstand and build in protection to the state trajectory. Nonetheless, operating reserve will continue to be a part of system operations for the foreseeable future. Thus, the presented formulations simulate the operator deployment of reserve.

Explicit formulations of rolling-horizon UC problems and ED problems that inherit from UC (including reserve committed by UC) are currently confined to ISO business practice manuals and are not accompanied by experimental results. Therefore, the formulations presented here contribute to existing UC literature by providing a framework for conducting rolling-horizon UC studies that allow for greater insight into actual system response to contingencies. Further, they contribute to the decentralized UC literature by demonstrating why and how conducting UC with faster periodicity is beneficial.

To improve the fidelity of the RHUC and ED models further, a few things should be done. The addition of "notification" times, or the amount of advance time that some generators need between being commanded and actually responding, can have implications for the responsiveness to faster UC. A generator with a minimum notification time of 20 minutes will not provide much additional flexibility if system UC is

conducted on a 20-minute basis as opposed to a 30-minute basis. This is challenging from an academic approach because very few, if any, realistic datasets include data for notification time. Further, the implementation of more detailed models for combined-cycle units is necessary if testing with realistic market data. Since the faster UC will be most often modulating the fast-responding combined-cycle units to recover from contingencies, an accurate assessment of those units' capabilities and costs are particularly important. Finally, a historical record of time-series forecasts in addition to simple historical wind data would help in understanding the vulnerabilities to wind intermittency.

6.5 Closing Statements

Electric power grid are trending more towards decentralization of decision-making and control. Even where larger, more centralized structures are forming, the need for distributed computation and scalability is being recognized. Although future decentralized power grids may offer radically more consumer choice, selection of electricity services, and accessibility, the benefits for today's computational challenges and operational scales are also apparent. Unit commitment is near the top of the list of power system-related computational challenges, and here it is shown that DUC may be a viable strategy to address it. The developments shown here are not representative of a commercial-ready solution for large-scale transmission grids, but the studies here show that it does hold legitimate potential. Collaborations with large-scale system operators can provide validation against many historical datasets and provide performance improvements based on their experience handling large UC problems. The critical nature of electric infrastructure means the path to online implementation is lengthy for electric grid operators even for long-proven technologies, but a long-term joint effort can provide rich results in this emerging field.

APPENDIX A

NOTATION

A.1 Abbreviations

ADMM	Alternating Direction Method of Multipliers
DUC	Decentralized Unit Commitment
ED	Economic Dispatch
GSF	Generation Shift Factor
ISO	Independent System Operator
MIP	Mixed Integer (Linear) Programming
MIQP	Mixed Integer Quadratic Programming
LODF	Line Outage Distribution Factor
NSR	Nonspinning Reserve
OPF	Optimal Power Flow
OR	Operating Reserve
QP	Quadratic Programming
R&F	Release and Fix
SR	Spinning Reserve
UC	Unit Commitment

A.2 General Nomenclature

A.2.1 Sets and Indices

$\nu \in \mathcal{P}$	Regions, running from 1 to n
$\nu' \in \Delta_i$	Set of regions (including region ν) connected to bus $i \in \mathcal{N}_\nu^{\mathrm{BB}}$

$g \in \mathcal{G}$	Generating units, running from 1 to G
$g \in \mathcal{G}_i$	Generating units in bus i
$i \in \mathcal{N}$	Set of buses, running from 1 to N
$i \in \mathcal{N}_\nu$	Set of all buses in region ν
$i \in \mathcal{N}_\nu^{\mathrm{BB}}$	Set of boundary buses in region ν
$i \in \mathcal{N}_\nu^{\mathrm{FB}}$	Set of foreign buses not in region ν connected to some bus in region ν
$i \in \mathcal{N}_\nu^{\mathrm{IB}}$	Set of internal buses in region ν
$ij \in \mathcal{E}$	Transmission lines, running from 1 to E
$ij \in \mathcal{E}'$	Set of contingency monitored transmission lines
$j \in \delta_i$	Set of all buses connected to bus i
$t \in \mathcal{T}$	Hourly periods, running from 1 to T hours
$\mathbf{x}_\nu \in S_\nu$	Set of all feasible points for region ν

A.2.2 Input Data and Other Parameters

α^m	Vector of primal residuals at iteration m
β^m	Vector of dual residuals at iteration m
$\gamma_{ij,k}$	Power flow sensitivity of line ij with respect to injection transfer from bus k to slack bus
ϵ^{Dual}	Dual tolerance
ϵ^{Pri}	Primal tolerance
ϵ^{Obj}	Relative objective function tolerance
$\lambda^m_{\nu,i,t}$	Value of dual variable corresponding to deviation from $\bar{\bar{\theta}}^m_{\nu,i,t}$, iteration m
$\mu^m_{\nu,i,t}$	Value of dual variable corresponding to deviation from $\bar{\bar{F}}^m_{\nu,ij,t}$, iteration m

ρ	Penalty factor
$\tilde{\theta}^m_{\nu,i,t}$	Optimal value of $\tilde{\theta}_{\nu,i,t}$ at iteration m
$\bar{\tilde{\theta}}^m_{\nu,i,t}$	Average value for $\tilde{\theta}_{\nu,i,t}$ at iteration m
$\omega^{q,m}_{\nu,t}$	Value of dual variable corresponding to deviation from $r^{q,m}_{\nu,t}$
B_{ij}	Element ij of DC power flow Jacobian
$C_\nu(.)$	Total cost function of region ν
C^{CS}_g	Cold startup cost of unit g (\$)
C^{HS}_g	Hot startup cost of unit g (\$)
C^{LV}_g	Linear energy cost of unit g (\$/MWh)
C^{NL}_g	No-load cost of unit g (\$/h)
C^{Q}_g	Quadratic energy cost of unit g (\$/MW^{2}h)
C^{SD}_g	Shutdown cost of unit g (\$)
CR(.)	Continuous relaxation of a discrete set
$d_{i,t}$	Expected load at bus i, time t
$\overline{F}_{ij}$	Active power flow limit of branch ij
$\tilde{F}^m_{\nu,ij,t}$	Optimal value of $\tilde{F}_{\nu,ij,t}$ at iteration m
$\bar{\tilde{F}}^m_{\nu,ij,t}$	Average value for $\tilde{F}_{\nu,ij,t}$ at iteration m
$\mathcal{L}_{\rho,\nu}(.)$	Augmented Lagrangian of region ν
$\overline{P}_g$	Maximum power output of unit g (MW)
$\underline{P}_g$	Minimum power output of unit g (MW)
$r^{q,m}_{\nu,t}$	Optimal value of $r^q_{\nu,t}$ at iteration m
$\bar{r}^{q,m}_{\nu,t}$	Target value for $r^q_{\nu,t}$ at iteration m
$\hat{r}^q_t$	System reserve requirement for product q (MW)

RD_g	Ramp-down rate of unit g (MW/h)
RU_g	Ramp-up rate of unit g (MW/h)
SD_g	Shutdown capability of unit g (MW)
SU_g	Startup capability of unit g (MW)
T_g^{CS}	Cold startup time of unit g (h)
T_g^{Init}	Number of hours unit g has been online (+) or offline (-) prior to the first period of the UC
TD_g	Minimum downtime of unit g (h)
TU_g	Minimum uptime of unit g (h)
t_{RR}	Computational time spent solving the root relaxation node for centralized cases (s)
$t_{1\%}$	Computational time spent finding a solution with 1% relative optimality gap (s)
t_{best}	Computational time spent finding the best binary-feasible solution (s)

A.2.3 Decision variables

$\theta_{i,t}$	Voltage phase angle at bus i, time t
$\tilde{\theta}_{\nu,i,t}$	Perception by region ν of voltage phase angle at bus $i \in \mathcal{N}_\nu^{\mathrm{BB}} \cup \mathcal{N}_\nu^{\mathrm{FB}}$, time t
$F_{ij,t}$	Power flow from bus i to bus j at time t
$\tilde{F}_{\nu,ij,t}$	Perception by region ν of power flow from bus $i \in \mathcal{N}_\nu^{\mathrm{BB}}$ to bus $j \in \mathcal{N}_\nu^{\mathrm{FB}}$ at time t
$P_{i,t}^{\mathrm{net}}$	Total real power injection from generators and loads at bus i, time t

p_{gt} Power output of unit g at hour t above the minimum output $\underline{P}_g$ (MW)

$r^q_{\nu,t}$ Amount of reserve product q provided by region ν at time t (MW)

r^q_{gt} Amount of reserve product q allocated to unit g at time t (MW)

u_{gt} Commitment status of unit g at hour t, equal to 1 if the unit is online and 0 if offline

$\mathbf{u}_\nu$ Vector of binary decision variables in region ν, including u_{gt} variables for all $g \in \mathcal{G}_\nu$ and $t \in \mathcal{T}$

v_{gt} Startup status of unit g, which takes the value of 1 if g starts up in hour t and 0 otherwise

v^{HS}_{gt} Hot startup status of unit g, which takes the value of 1 in the hour t if the unit starts up and shuts down in the interval $[t - T^{\text{CS}}_g - TD_g, t-1]$

w_{gt} Shutdown status of unit g, which takes the value of 1 if g shuts down in hour t and 0 otherwise

$\mathbf{x}_\nu$ Vector of all decision variables in region ν, including variables u_{gt}, v_{gt}, v^{HS}_{gt}, w_{gt}, p_{gt}, and r^q_{gt} of generators $g \in \mathcal{G}_\nu$, variables $\theta_{i,t}(.)$ for internal buses $i \in \mathcal{N}^{\text{IB}}_\nu$, and $\tilde{\theta}_{\nu,i,t}$ of all buses $i \in \mathcal{N}^{\text{BB}}_\nu \cup \mathcal{N}^{\text{FB}}_\nu$, for all $t \in \mathcal{T}$

$\mathbf{y}_\nu$ Vector of continuous decision variables in region ν, including all entries of $\mathbf{x}_\nu$ not present in $\mathbf{u}_\nu$

A.3 Nomenclature for RHUC Formulation

A.3.1 Sets and Indices

$\nu \in \mathcal{P}$	Regions, running from 1 to n
$\nu' \in \Delta_i$	Set of regions (including region ν) connected to bus $i \in \mathcal{N}_\nu^{\mathrm{BB}}$
$g \in \mathcal{G}$	Generating units, running from 1 to G
$g \in \mathcal{G}_i$	Generating units in bus i
$g \in \mathcal{G}^{\mathrm{SS}}$	Slow-start generating units g
$g \in \mathcal{G}^{\mathrm{FS}}$	Fast-start generating units g
$g \in \mathcal{G}^{\mathrm{SSh}}$	Slow-stop generating units g
$g \in \mathcal{G}^{\mathrm{FSh}}$	Fast-stop generating units g
$g \in \mathcal{G}^{\mathrm{R}}$	Renewable generating units g
$i \in \mathcal{N}$	Set of buses, running from 1 to N
$i \in \mathcal{N}_\nu$	Set of all buses in region ν
$i \in \mathcal{N}_\nu^{\mathrm{BB}}$	Set of boundary buses in region ν
$i \in \mathcal{N}_\nu^{\mathrm{FB}}$	Set of foreign buses not in region ν connected to some bus in region ν
$i \in \mathcal{N}_\nu^{\mathrm{IB}}$	Set of internal buses in region ν
$ij \in \mathcal{E}$	Transmission lines, running from 1 to E
$ij \in \mathcal{E}'$	Set of contingency monitored transmission lines
$ij \in \mathcal{E}_\nu^{\mathrm{B}}$	Set of boundary lines of prosumer ν
$ij \in \mathcal{E}_\nu^{\mathrm{I}}$	Set of internal lines of prosumer ν
$j \in \delta_i$	Set of all buses connected to bus i
$k \in \mathcal{L}$	Set of all loads, running from 1 to L
$t \in \mathcal{T}$	Hourly periods from master UC, running from 1 to T hours
$t \in \mathcal{T}_h$	Time periods in rolling UC, running from t_1 to t_h
$\mathbf{x}_\nu \in S_\nu$	Set of all feasible points for region ν

A.3.2 Input Data and Other Parameters

B_{ij}	Element ij of DC power flow Jacobian
$C(\mathbf{x}, t)$	Total cost function with state vector $\mathbf{x}$ at time t
C_g^{CS}	Cold startup cost of unit g (\$)
C_g^{HS}	Hot startup cost of unit g (\$)
C_g^{LV}	Linear energy cost of unit g (\$/MWh)
C_g^{NL}	No-load cost of unit g (\$/h)
C_g^{Q}	Quadratic energy cost of unit g (\$/MW^{2}h)
C_g^{SD}	Shutdown cost of unit g (\$)
$d_{k,t}^{\mathrm{F}}$	Expected fixed demand of load k, time t
$d_k^{\max}$	Maximum curtailable demand of load k
$\overline{F}_{ij}$	Active power flow limit of branch ij
$\tilde{F}_{\nu,ij,t}^m$	Optimal value of $\tilde{F}_{\nu,ij,t}$ at iteration m
$\bar{\tilde{F}}_{\nu,ij,t}^m$	Average value for $\tilde{F}_{\nu,ij,t}$ at iteration m
$p_{i,t}^{\mathrm{Rmax}}$	Maximum forecast renewable capacity at bus i, time t
p_g^0	Initial power output of unit g (MW)
$\overline{P}_g$	Maximum power output of unit g (MW)
$\underline{P}_g$	Minimum economic power output of unit g (MW)
$\hat{r}_t^q$	Reserve product q requirement at time t
RD_g	Ramp-down rate of unit g (MW/h)
RU_g	Ramp-up rate of unit g (MW/h)
SD_g	Shutdown capability of unit g (MW)
SU_g	Startup capability of unit g (MW)
T_g^{CS}	Cold startup time of unit g (h)

T_g^{Init} Number of hours unit g has been online (+) or offline (-) prior to the first period of the UC

TD_g Minimum downtime of unit g (h)

TD_g^0 Initial minimum downtime of unit g (h)

TU_g Minimum uptime of unit g (h)

TU_g^0 Initial minimum uptime of unit g (h)

t_{int} Time length of UC solution interval (h)

t_{FS} Startup time of fast-start generators (assume 20 min usually)

u_{gt}^* Predetermined commitment status of unit g at time t, equal to 1 if the unit is online and 0 if offline

$\hat{u}_{gt}$ Suggested commitment status of unit g at time t inherited by 24-hour UC

v_{gt}^* Predetermined startup status of unit g at time t, equal to 1 if the unit is in startup and 0 otherwise

w_{gt}^* Predetermined shutdown status of unit g at time t, equal to 1 if the unit is in shutdown and 0 otherwise

A.3.3 Decision variables

$\theta_{i,t}$ Voltage phase angle at bus i, time t

d_{kt}^{C} Controllable part of load k served at time t

$F_{ij,t}$ Power flow from bus i to bus j at time t

$P_{i,t}^{\text{net}}$ Total real power injection from generators and loads at bus i, time t

p_{gt} Power output of unit g at hour t above $\underline{P}_g$ (MW)

r_{gt}^q Amount of reserve product q scheduled by unit g at time t (MW)

u_{gt}	Commitment status of unit g at time t, equal to 1 if the unit is online and 0 if offline
$\mathbf{u}_\nu$	Vector of binary decision variables in region ν, including u_{gt} variables for all $g \in \mathcal{G}_\nu$ and $t \in \mathcal{T}$
v_{gt}	Startup status of unit g, which takes the value of 1 if g starts up at time t and 0 otherwise
v_{gt}^{HS}	Hot startup status of unit g, which takes the value of 1 at time t if the unit starts up and shuts down in the interval $[t - T_g^{\mathrm{CS}} - TD_g, t-1]$
$\tilde{v}_{gt}$	Transitional startup status of unit g, a rational value between 0 and 1
w_{gt}	Shutdown status of unit g, which takes the value of 1 if g shuts down at time t and 0 otherwise
$\tilde{w}_{gt}$	Transitional shutdown status of unit g, a rational value between 0 and 1
$\mathbf{x}_g$	Vector of all decision variables for generator g
$\mathbf{x}_k$	Vector of all decision variables for load k

A.4 Nomenclature for ED Formulation

A.4.1 Sets and Indices

$\nu \in \mathcal{P}$	Regions, running from 1 to n
$\nu' \in \Delta_i$	Set of regions (including region ν) connected to bus $i \in \mathcal{N}_\nu^{\mathrm{BB}}$
$g \in \mathcal{G}$	Generating units, running from 1 to G
$g \in \mathcal{G}_i$	Generating units in bus i
$g \in \mathcal{G}^{\mathrm{SS}}$	Slow-start generating units g
$g \in \mathcal{G}^{\mathrm{FS}}$	Fast-start generating units g
$g \in \mathcal{G}^{\mathrm{SSh}}$	Slow-stop generating units g
$g \in \mathcal{G}^{\mathrm{FSh}}$	Fast-stop generating units g
$g \in \mathcal{G}^{\mathrm{R}}$	Renewable generating units g
$i \in \mathcal{N}$	Set of buses, running from 1 to N
$i \in \mathcal{N}_\nu$	Set of all buses in region ν
$i \in \mathcal{N}_\nu^{\mathrm{BB}}$	Set of boundary buses in region ν
$i \in \mathcal{N}_\nu^{\mathrm{FB}}$	Set of foreign buses not in region ν connected to some bus in region ν
$i \in \mathcal{N}_\nu^{\mathrm{IB}}$	Set of internal buses in region ν
$ij \in \mathcal{E}$	Transmission lines, running from 1 to E
$ij \in \mathcal{E}'$	Set of branch outages included in contingency constraints
$j \in \delta_i$	Set of all buses connected to bus i
$k \in \mathcal{L}$	Set of all loads, running from 1 to L
$t \in \mathcal{T}$	Hourly periods, running from 1 to T hours
$\mathbf{x}_\nu \in S_\nu$	Set of all feasible points for region ν

A.4.2 Input Data and Other Parameters

B_{ij}	Element ij of DC power flow Jacobian
$C(\mathbf{x}, t)$	Total cost function with state vector $\mathbf{x}$ at time t
C_g^{LV}	Linear energy cost of unit g (\$/MWh)
C_g^{NL}	No-load cost of unit g (\$/h)
C_g^{Q}	Quadratic energy cost of unit g (\$/MW^{2}h)
$d_{k,t}^{\mathrm{F}}$	Expected fixed demand of load k, time t
$d_k^{\max}$	Maximum curtailable demand of load k
$\overline{F}_{ij}$	Active power flow limit of branch ij
$\tilde{F}_{\nu,ij,t}^m$	Optimal value of $\tilde{F}_{\nu,ij,t}$ at iteration m
$\bar{\tilde{F}}_{\nu,ij,t}^m$	Average value for $\tilde{F}_{\nu,ij,t}$ at iteration m
$\overline{P}_g$	Maximum power output of unit g (MW)
$p_{i,t}^{\mathrm{Rmax}}$	Maximum forecast renewable capacity at bus i, time t
$\underline{P}_g$	Minimum economic power output of unit g (MW)
$\hat{r}_t^q$	Total system reserve product q scheduled by UC at time t
RD_g	Ramp-down rate of unit g (MW/h)
RU_g	Ramp-up rate of unit g (MW/h)
SD_g	Shutdown capability of unit g (MW)
SU_g	Startup capability of unit g (MW)
T_g^{CS}	Cold startup time of unit g (h)
T_g^{Init}	Number of hours unit g has been online (+) or offline (-) prior to the first period of the UC
TD_g	Minimum downtime of unit g (h)
TU_g	Minimum uptime of unit g (h)

t^-	Time of the most recent chronological UC decision
t^+	Time of the next chronological UC decision
t_{int}	Time length of ED interval, usually equivalent to 10 min (h)
t_p	Time period of prior ED interval
u_{gt}	Commitment status of unit g at time t, equal to 1 if the unit is online and 0 if offline
v_{gt}	Startup status of unit g at time t, equal to 1 if the unit is in startup and 0 otherwise
w_{gt}	Shutdown status of unit g at time t, equal to 1 if the unit is in shutdown and 0 otherwise

A.4.3 Decision variables

$\theta_{i,t}$	Voltage phase angle at bus i, time t
d_{kt}^{C}	Controllable part of load k served at time t
$F_{ij,t}$	Power flow from bus i to bus j at time t
$P_{i,t}^{\text{net}}$	Total real power injection from generators and loads at bus i, time t
p_{gt}	Power output of unit g at hour t (MW)
r_{gt}^{q}	Amount of reserve product q scheduled by unit g at time t (MW)
$\mathbf{x}_g$	Vector of all decision variables for generator g

REFERENCES

[1] Alemany, J., Moitre, D., and Magnago, F., "Benders decomposition applied to security constrained unit commitment," *IEEE Latin America Transactions*, vol. 11, no. 1, pp. 421–425, 2013.

[2] Ardakani, A. and Bouffard, F., "Identification of umbrella constraints in dc-based security-constrained optimal power flow," *IEEE Trans. Power Systems*, vol. 28, pp. 3924–3934, Nov 2013.

[3] Bakirtzis, A. G. and Biskas, P. N., "A decentralized solution to the DC-OPF of interconnected power systems," *IEEE Trans. Power Systems*, vol. 18, no. 3, pp. 1007–1013, 2003.

[4] Batut, J. and Renaud, A., "Daily generation scheduling optimization with transmission constraints: a new class of algorithms," *IEEE Trans. Power Systems*, vol. 7, no. 3, pp. 982–989, 1992.

[5] Biskas, P. and Bakirtzis, A., "Decentralised security constrained dc-opf of interconnected power systems," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 151, pp. 747–754, Nov 2004.

[6] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[7] Brandwajn, V., "Efficient bounding method for linear contingency analysis," *IEEE Trans. Power Systems*, vol. 3, pp. 38–43, Feb 1988.

[8] Brower, M., "Wind forecasting and integration from minutes, to hours, to days," in *MIT Wind Week Integration Workshop*, January 2011.

[9] Capitanescu, F. and Wehenkel, L., "A new iterative approach to the corrective security-constrained optimal power flow problem," *IEEE Trans. Power Systems*, vol. 23, pp. 1533–1541, Nov 2008.

[10] Carrion, M. and Arroyo, J. M., "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Trans. Power Systems*, vol. 21, no. 3, pp. 1371–1378, 2006.

[11] Chakrabarti, S., Kraning, M., Chu, E., Baldick, R., and Boyd, S., "Security constrained optimal power flow via proximal message passing," in *Power Systems Conference (PSC), 2014 Clemson University*, pp. 1–8, March 2014.

[12] Chen, Y., Wang, Q., Wang, X., and Guan, Y., "Applying robust optimization to miso look-ahead commitment," in *PES General Meeting | Conference Exposition, 2014 IEEE*, pp. 1–5, July 2014.

[13] Chen, Y., Wang, X., and Wang, Q., "Overcoming computational challenges on large-scale security constrained unit commitment (scuc) problems," in *FERC Technical Conference on increasing Real-Time and Day-Ahead Market Efficiency through Improved Software, June 23-25, 2014*.

[14] Chung, K. H., Kim, B. H., and Hur, D., "Multi-area generation scheduling algorithm with regionally distributed optimal power flow using alternating direction method," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 9, pp. 1527–1535, 2011.

[15] Clement-Nyns, K., Haesen, E., and Driesen, J., "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *IEEE Trans. Power Systems*, vol. 25, pp. 371–380, Feb 2010.

[16] Constantinescu, E., Zavala, V., Rocklin, M., Lee, S., and Anitescu, M., "A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation," *IEEE Trans. Power Systems*, vol. 26, pp. 431–441, Feb 2011.

[17] Correia, P., "Decentralised unit commitment in a market structure: problem formulation and solution advancement," *IEE Proceedings-Generation, Transmission and Distribution*, vol. 153, no. 1, pp. 121–126, 2006.

[18] Costley, M. and Feizollahi, M. J., "RHUC test cases." `http://www.ece.gatech.edu/research/labs/aces/pages/rhuctestdata.zip`, 2015.

[19] DoE, "Department of energy smart grid system report," 2009.

[20] Dong, F., Xu, X., and Zhang, X., "Parallel contingency analysis solution based on openmp," in *North American Power Symposium (NAPS), 2014*, pp. 1–4, Sept 2014.

[21] Dzung, P. and Kalagnanam, J., "Distributed methods for solving the security-constrained optimal power flow problem," in *Innovative Smart Grid Technologies (ISGT), 2012*, pp. 1–7.

[22] Ejebe, G. and Wollenberg, B., "Automatic contingency selection," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-98, pp. 97–109, Jan 1979.

[23] EnerNex Corporation, "Eastern wind integration and transmission study." `http://www.nrel.gov/docs/fy11osti/47078.pdf`, 2011.

[24] EPRI, "Smart grid enterprise architecture interest group," 2013.

[25] Feinberg, E., Hu, J., and Yuan, E., "A parallel rolling horizon scheme for large scale security constrained unit commitment problems with wind power generation," in *Proceedings of ENERGY 2012*.

[26] Feizollahi, M. J. and Costley, M., "DUC test cases." `http://www.ece.gatech.edu/research/labs/aces/pages/ductestdata.zip`, 2014.

[27] Feizollahi, M. J., Costley, M., Ahmed, S., and Grijalva, S., "Large-scale decentralized unit commitment," *International Journal of Electrical Power & Energy Systems, under review*, 2014.

[28] Finardi, E. C. and da Silva, E. L., "Solving the hydro unit commitment problem via dual decomposition and sequential quadratic programming," *IEEE Trans. Power Systems*, vol. 21, no. 2, pp. 835–844, 2006.

[29] Fischer, R. and Serra, P., "Energy prices in the presence of plant indivisibilities," *Energy Economics*, vol. 25, no. 4, pp. 303–314, 2003.

[30] Fu, Y., Shahidehpour, M., and Li, Z., "Long-term security-constrained unit commitment: hybrid Dantzig-Wolfe decomposition and subgradient approach," *IEEE Trans. Power Systems*, vol. 20, no. 4, pp. 2093–2106, 2005.

[31] Galiana, F. D., "Bound estimates of the severity of line outages in power system contingency analysis and ranking," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-103, no. 9, pp. 2612–2624, 1984.

[32] Gatterbauer, W., "Economic efficiency of decentralized unit commitment from a generator's perspective," *Engineering Electricity Services of the Future*, 2010.

[33] Granada E, M., Rider, M. J., Mantovani, J. R. S., and Shahidehpour, M., "Decentralized AC power flow for real-time multi-TSO power system operation," in *Power and Energy Society General Meeting, 2010*, pp. 1–7.

[34] Grijalva, S. and Costley, M., "Decentralized energy management architecture for service-oriented cyber-physical electric grids." 2013.

[35] Grijalva, S., Costley, M., and Ainsworth, N., "Prosumer-based control architecture for the future electricity grid," in *IEEE Multi-conference on Systems and Control*.

[36] Grijalva, S. and Tariq, M. U., "Prosumer-based smart grid architecture enables a flat, sustainable electricity industry," in *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*, pp. 1–6.

[37] Guo, J., Fu, Y., Li, Z., and Shahidehpour, M., "Direct calculation of line outage distribution factors," *IEEE Trans. Power Systems*, vol. 24, pp. 1633–1634, Aug 2009.

[38] Heredia, F. J., Rider, M. J., and Corchero, C., "Optimal bidding strategies for thermal and combined cycle units in the day-ahead electricity market with bilateral contracts," in *Power & Energy Society General Meeting, 2009*, pp. 1–6.

[39] Hodge, B. and Milligan, M., "Wind power forecasting error distributions over multiple timescales," in *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–8, July 2011.

[40] Honarvar Nazari, M., Costello, Z., Feizollahi, M. J., Grijalva, S., and Egerstedt, M., "Distributed frequency control of prosumer-based electric energy systems," *IEEE Trans. Power Systems, in press*, 2014.

[41] Huang, Z., Chen, Y., and Nieplocha, J., "Massive contingency analysis with high performance computing," in *Power Energy Society General Meeting, 2009. PES '09. IEEE*, pp. 1–8, July 2009.

[42] Hui, H., Yu, C.-N., and Moorty, S., "Reliability unit commitment in the new ERCOT nodal electricity market," in *Power & Energy Society General Meeting, 2009*, pp. 1–8.

[43] Jabr, R. A., "Tight polyhedral approximation for mixed-integer linear programming unit commitment formulations," *Generation, Transmission & Distribution, IET*, vol. 6, no. 11, pp. 1104–1111, 2012.

[44] Kargarian, A., Fu, Y., and Li, Z., "Distributed security-constrained unit commitment for large-scale power systems," *IEEE Trans. Power Systems*, vol. PP, no. 99, pp. 1–12, 2014.

[45] Kargarian, A. and Fu, Y., "System of systems based security-constrained unit commitment incorporating active distribution grids," *IEEE Trans. Power Systems*, vol. 29, pp. 2489–2498, Sept 2014.

[46] Khaitan, S. and McCalley, J., "Empower: An efficient load balancing approach for massive dynamic contingency analysis in power systems," in *High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion:*, pp. 289–298, Nov 2012.

[47] Kim, B. H. and Baldick, R., "Coarse-grained distributed optimal power flow," *IEEE Trans. Power Systems*, vol. 12, no. 2, pp. 932–939, 1997.

[48] Kim, B. H. and Baldick, R., "A comparison of distributed optimal power flow algorithms," *IEEE Trans. Power Systems*, vol. 15, no. 2, pp. 599–604, 2000.

[49] Kim, S.-J. and Giannakis, G., "Scalable and robust demand response with mixed-integer constraints," *Smart Grid, IEEE Transactions on*, vol. 4, pp. 2089–2099, Dec 2013.

[50] Li, C., Svoboda, A. J., Xiaohong, G., and Singh, H., "Revenue adequate bidding strategies in competitive electricity markets," *IEEE Trans. Power Systems*, vol. 14, no. 2, pp. 492–497, 1999.

[51] Li, M. and Luh, P. B., "A decentralized framework of unit commitment for future power markets," in *Power and Energy Society General Meeting, 2013*, pp. 1–5.

[52] Li, T. and Shahidehpour, M., "Price-based unit commitment: a case of lagrangian relaxation versus mixed integer programming," *IEEE Trans. Power Systems*, vol. 20, no. 4, pp. 2015–2025, 2005.

[53] Ma, H. and Shahidehpour, S. M., "Transmission constrained unit commitment based on Benders decomposition," *International Journal of Electrical Power & Energy Systems*, vol. 20, no. 4, pp. 287–94, 1998.

[54] Maheswari, S. and Vijayalakshmi, C., "A lagrangian decomposition model for unit commitment problem," *International Journal of Computer Applications*, vol. 43, no. 12, pp. 21–25, 2012.

[55] Mashhour, E. and Moghaddas-Tafreshi, S. M., "Bidding strategy of virtual power plant for participating in energy and spinning reserve markets - part i: Problem formulation," *IEEE Trans. Power Systems*, vol. 26, no. 2, pp. 949–956, 2011.

[56] Monticelli, A., Pereira, M., and Granville, S., "Security-constrained optimal power flow with post-contingency corrective rescheduling," *IEEE Trans. Power Systems*, vol. 2, pp. 175–180, Feb 1987.

[57] Morales-Espana, G., Latorre, J. M., and Ramos, A., "Tight and compact MILP formulation for the thermal unit commitment problem," *IEEE Trans. Power Systems*, vol. 28, no. 4, pp. 4897–4908, 2013.

[58] Moreira, A., Street, A., and Arroyo, J., "An adjustable robust optimization approach for contingency-constrained transmission expansion planning," *IEEE Trans. Power Systems*, vol. PP, no. 99, pp. 1–10, 2014.

[59] Motto, A. L. and Galiana, F. D., "Equilibrium of auction markets with unit commitment: The need for augmented pricing," *IEEE Trans. Power Systems*, vol. 17, no. 3, pp. 798–805, 2002.

[60] Motto, A. L. and Galiana, F. D., "Coordination in markets with nonconvexities as a mathematical program with equilibrium constraints-part i: a solution procedure," *IEEE Trans. Power Systems*, vol. 19, no. 1, pp. 308–316, 2004.

[61] NERC, "NERC IVGTF task 2.4 report." `http://www.nerc.com/files/ivgtf2-4.pdf`, 2011.

[62] Nogales, F., Prieto, F., and Conejo, A., "A decomposition methodology applied to the multi-area optimal power flow problem," *Annals of Operations Research*, vol. 120, no. 1-4, pp. 99–116, 2003.

[63] Ostrowski, J., Anjos, M. F., and Vannelli, A., "Tight mixed integer linear programming formulations for the unit commitment problem," *IEEE Trans. Power Systems*, vol. 27, no. 1, pp. 39–46, 2012.

[64] Ott, A., "Unit commitment in the PJM day-ahead and real-time markets," in *FERC Technical Conference on increasing Market and Planning Efficiency through Improved Software, June, 2010.*

[65] Pedrasa, M., Spooner, T., and MacGill, I., "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *Smart Grid, IEEE Transactions on*, vol. 1, pp. 134–143, Sept 2010.

[66] PSERC, "The future grid to enable sustainable energy systems," 2013.

[67] Rajan, D. and Takriti, S., "Minimum up/down polytopes of the unit commitment problem with start-up costs." IBM Report, 2005.

[68] Ramachandran, T., Costello, Z., Kingston, P., Egerstedt, M., and Grijalva, S., "Distributed power allocation in prosumer networks," in *IFAC Workshop on Estimation and Control of Networked Systems.*

[69] Shrestha, G. B., Pokharel, B. K., Lie, T. T., and Fleten, S. E., "Price-based unit commitment for bidding under price uncertainty," *Generation, Transmission & Distribution, IET*, vol. 1, no. 4, pp. 663–669, 2007.

[70] Simoglou, C. K., Biskas, P. N., and Bakirtzis, A. G., "Optimal self-scheduling of a thermal producer in short-term electricity markets by MILP," *IEEE Trans. Power Systems*, vol. 25, no. 4, pp. 1965–1977, 2010.

[71] Sioshansi, R., O'Neill, R., and Oren, S. S., "Economic consequences of alternative solution methods for centralized unit commitment in day-ahead electricity markets," *IEEE Trans. Power Systems*, vol. 23, no. 2, pp. 344–352, 2008.

[72] Sioshansi, R., Oren, S., and O'Neill, R., "The cost of anarchy in self-commitment based electricity markets," *Competitive Electricity Markets: Design, Implementation and Performance*, 2008.

[73] Stefopoulos, G., Yang, F., Cokkinides, G., and Meliopoulos, A., "Advanced contingency selection methodology," in *Power Symposium, 2005. Proceedings of the 37th Annual North American*, pp. 67–73, Oct 2005.

[74] Thomas, J. and Grijalva, S., "Flexible security-constrained optimal power flow," *IEEE Trans. Power Systems*, vol. PP, no. 99, pp. 1–8, 2014.

[75] Tuohy, A., Denny, E., and O'Malley, M., "Rolling unit commitment for systems with significant installed wind capacity," in *Power Tech, 2007 IEEE Lausanne*, pp. 1380–1385.

[76] Wang, Q., Watson, J., and Guan, Y., "Two-stage robust optimization for n-k contingency-constrained unit commitment," *IEEE Trans. Power Systems*, vol. 28, pp. 2366–2375, Aug 2013.

[77] Wood, A. J. and Wollenberg, B. F., *Power Generation, Operation, and Control*. Wiley, 2 ed., 1996.

[78] Xu, J. and Christie, R. D., "Decentralized unit commitment in competitive energy markets," *International Series in Operations Research and Management Science*, pp. 293–316, 2001.

[79] Yamin, H. Y. and Shahidehpour, S. M., "Self-scheduling and energy bidding in competitive electricity markets," *Electric Power Systems Research*, vol. 71, no. 3, pp. 203–209, 2004.

[80] Yong, F., Shahidehpour, M., and Li, Z., "Security-constrained unit commitment with ac constraints," *IEEE Trans. Power Systems*, vol. 20, no. 2, pp. 1001–1013, 2005.

[81] Zhao, F., Litvinov, E., and Tongxin, Z., "A marginal equivalent decomposition method and its application to multi-area optimal power flow problems," *IEEE Trans. Power Systems*, vol. 29, no. 1, pp. 53–61, 2014.

[82] Ziadi, Z., Taira, S., Oshiro, M., and Funabashi, T., "Optimal power scheduling for smart grids considering controllable loads and high penetration of photovoltaic generation," *Smart Grid, IEEE Transactions on*, vol. 5, pp. 2350–2359, Sept 2014.

[83] Zimmerman, R. D., Murillo-Sánchez, C. E., and Gan, D., "A MATLAB power system simulation package," 2005.

VITA

Mitcham Costley (BSEE 2008, MSECE 2009) is a PhD Candidate in the Advanced Computational Electricity Systems (ACES) Laboratory at Georgia Tech. As an undergraduate, he gained industry-side experience in electrical power grids at the consulting engineering and utility levels. He designed, built, and tested RF circuits and embedded systems while pursuing a master's degree. Mitch also constructed and developed control software for a lab-scale testbed for microgrid voltage and frequency control. In his PhD work, he is working on decentralized control and operations, especially unit commitment, for future power grids as part of the ARPA-E GENI program. Mitch is experienced in translating complex mathematical algorithms for solving these problems into efficient software on both conventional machines and large-scale parallel architectures. He has broad exposure to power systems problems across many layers, from electricity markets and system control down to device design and control.